A NEW CLASS OF UNBALANCED CAST CIPHERS AND ITS SECURITY ANALYSIS

XIA ZHU

A New Class of Unbalanced CAST Ciphers and Its Security Analysis

by

©Xia Zhu

A thesis submitted to the

School of Graduate Studies

in partial fulfillment of the requirements for

the degree of Master of Engineering

Faculty of Engineering and Applied Science

Memorial University of Newfoundland

April, 1997

St. John's Newfoundland Canada

Abstract

The original CAST cipher is an efficient and secure private-key block cipher designed to be an alternative to the Data Encryption Standard (DES). In this thesis, we present a new class of unbalanced CAST ciphers which employs the same structure of S-box and round function of the original CAST cipher but has a lower memory requirement. Unbalanced CAST ciphers with one or two 8×32 S-boxes in the round function require only 1/4 or 1/2 the memory of the original CAST cipher, respectively.

This thesis examines the application of differential and linear cryptanalysis, two of the most powerful methodologies for attacking private-key block ciphers, to the unbalanced CAST ciphers. The results of analysis show that a 48-round unbalanced CAST cipher with one 8×32 S-box and a 24-round unbalanced CAST cipher with two 8×32 S-boxes, both of which are equivalent to a 12-round original CAST cipher in efficiency, are resistant to both differential and linear cryptanalysis.

We also investigate the unbalanced CAST ciphers from the perspective of information theory. The results suggest that the maximum static and dynamic input-output bit information leakages for the unbalanced CAST ciphers constructed by 8×32 S-boxes are much smaller than for DES.

The conclusion reached by the thesis is that unbalanced CAST ciphers can be considered to be efficient, secure ciphers which require less memory than the original CAST cipher.

Dedicated to my wife Boquan Xie.

For her encouragement and support, I am eternally grateful.

Acknowledgments

I would like to especially thank my supervisor, Dr. Howard Heys, for his research guidance, financial support and personal encouragement throughout this thesis.

Contents

List of Figures

List of Tables

List of Symbols

N	Cipher block size
R	Number of rounds in a cipher
K	Master key block used in a cipher
K_i	Subkey block used in i-th round
L_i	Left half block in i-th round
R_i	Right half block in i-th round
m	Number of input bits of an S-box
n	Number of output bits of an S-box
P	Plaintext block
C	Ciphertext block
P_{Ω_r}	Probability of an r-round characteristic
N_D	Number of chosen plaintexts required in differential cryptanalysis
P_L	Probability of a linear approximation
N_L	Number of known plaintexts required in linear cryptanalysis
M	Number of S-boxes in a round function

Chapter 1

Introduction

In recent years, with the rapid growth of computer networks, the threat of intercepting and/or modifying information during its transfer across a public communication channel has increased. *Cryptography*, the science of making information unintelligible and unmodifiable by an unauthorized interceptor and still accessible or verifiable by a legitimate receiver, is becoming more and more important in the field of communications. Figure 1.1 illustrates a general cryptographic system.

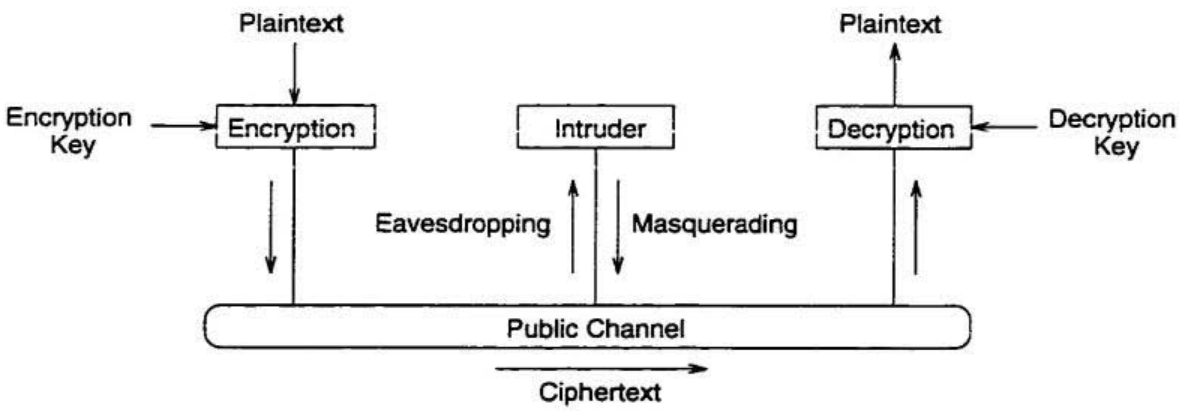

Figure 1.1: A General Cryptographic System

Encryption, performed by a transmitter, is a transforming process through which the original information is replaced by the secret information, while *decryption* is a

reverse process performed by a receiver. The message to be encrypted is referred to as the *plaintext* and the encrypted message is referred to as the *ciphertext*. A set of rules by which a transmitter encrypts the plaintext and a receiver decrypts the ciphertext is called a *cipher*. Normally the operation of the cipher depends on a *key*. Cryptography is the science of designing a cipher, while *cryptanalysis* is the art of breaking the cipher, that is, essentially determining the key. *Cryptology* includes both cryptography and cyptanalysis.

Cryptography may generally be divided into two categories: public-key ciphers and private-key ciphers. In a public-key cipher, the encryption key and the decryption key are different. Since it is computationally infeasible to determine a decryption key given an encryption key, the encryption key can be made public. Therefore, any transmitter can send an encrypted message with a public encryption key, while only the receiver with a secret decryption key can decrypt the ciphertext correctly. In a private-key cipher, the encryption key and the decryption key are the same and kept secret. The key must be distributed by a secure channel before any ciphertext is transmitted. In some cases, this may be difficult to achieve. However, generally private-key ciphers have a much higher encryption/decryption rate than public-key ciphers. In this thesis, we focus our attention on private-key ciphers which encrypt and decrypt data in blocks of bits.

1.1 Motivation for the Research

The most widely used private-key block encryption algorithm, the Data Encryption Standard (DES) [24] was first approved by the National Bureau of Standards (NBS)

in 1977 and was most recently reaffirmed in 1993, until December 1998. DES has been involved in controversy since its release. Its 56-bit key size has received wide criticism and its design principles are still unknown. After twenty years, DES is nearing the end of its useful life and is theoretically breakable by two powerful cryptanalytic attacks, differential and linear cryptanalysis [3, 19]. In addition, DES was explicitly designed for fast hardware implementation and has a slow software performance [21].

The National Institute of Standards and Technology (NIST) is initiating a process to develop a Federal Information Processing Standard (FIPS) for an Advanced Encryption Standard (AES) incorporating an Advanced Encryption Algorithm (AEA) as a replacement standard of DES at the 1998 review. As the first step in this process, draft minimum acceptability requirements and draft criteria to evaluate candidate algorithms have been issued:

- AES shall be publicly defined.
- AES shall be a symmetric private-key block cipher.
- AES shall be designed so that the key length may be increased as needed.
- AES shall be implementable in both hardware and software.
- AES shall be either freely available or available under terms consistent with the American National Standards Institute (ANSI) patent policy.
- Algorithms which meet the above requirements will be judged based on the following factors:
    * security,
    * computational efficiency,

* memory requirements,
* hardware and software suitability,
* simplicity,
* flexibility, and
* licensing requirements.

The original CAST cipher [2] is a symmetric block cipher and appears to be resistant to differential and linear cryptanalysis [18, 14]. It is easily implemented by software and has a good encryption/decryption performance on 32-bit microprocessors because of using four large 8×32 substitutions (S-boxes) and eliminating the need of permutations (P-boxes) which are awkward to implement in software. However, large S-boxes require more memory to store their lookup tables. This might be unacceptable in some implementations where the memory is extremely restricted.

We present a family of ciphers referred to as unbalanced CAST ciphers, which employ the same type of S-box and round function as the original CAST cipher and require a variable amount of memory depending on the chosen parameters.

1.2 Contributions of this Research

In this thesis, we present a new class of private-key block ciphers known as unbalanced CAST ciphers, incorporating the same structure of the S-box and round function of the original CAST cipher. The ciphers in this new class are simple, fast, and suitable for software and hardware implementations. They have a variable memory requirement and a variable number of rounds. We analyze the security of this new class of ciphers with respect to differential and linear cryptanalysis and its information

theoretic properties so that the user can explicitly manipulate the trade-off between higher speed and higher security. The results of the analysis show that the unbalanced CAST ciphers with proper parameters and an appropriate number of rounds are secure and promising ciphers.

1.3 Outline of Thesis

The organization of the remainder of this thesis is as following: Chapter 2 gives an overview of previous research that is directly relevant to our work. Chapter 3 presents a detailed discussion of differential and linear cryptanalysis techniques. Chapter 4 describes a new general class of unbalanced CAST ciphers and its design decisions. Chapter 5 examines the resistance of the unbalanced CAST ciphers with a set of typical parameters to differential cryptanalysis. Chapter 6 examines the resistance of the unbalanced CAST ciphers to linear cryptanalysis. Chapter 7 views the unbalanced CAST ciphers from the perspective of information theory. Finally, Chapter 8 provides a summary of the results of the thesis and some suggestions for further research.

Chapter 2

Review of Previous Research

In this chapter we present a review of previous research on private-key block ciphers which is directly related to our work.

2.1 Architectures

In this section, we introduce two main architectures which are widely quoted in designing private-key block ciphers. Other architectures are also detailed in the next section.

2.1.1 Substitution Permutation Networks

In 1949, Shannon [31] proposed two different concepts of confusion and diffusion in cryptography. Such concepts are embodied in a Substitution-Permutation Network (SPN), proposed by Feistel [11] and Feistel, et al [12].

The SPN consists of a number of substitution-permutation layers (SP layers), each of which is composed of several small sub-block substitutions (S-boxes) and a large bit permutation (P-box). The S-box which provides nonlinear input-output bit transformations fulfills Shannon's concept of confusion. The P-box which offers linear

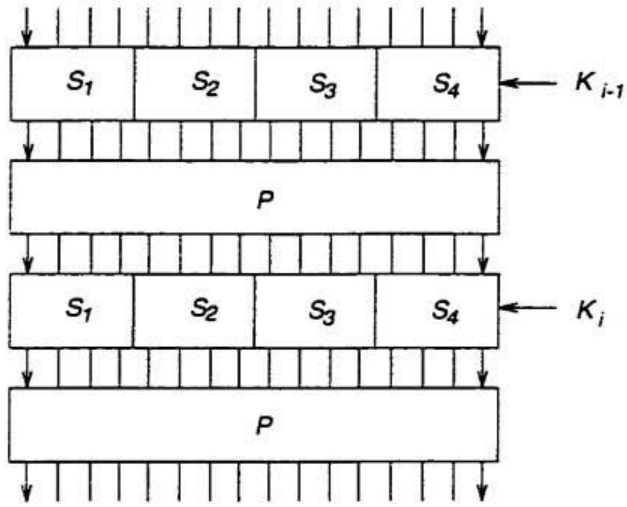

Figure 2.1: An Example of Substitution-Permutation Networks

bit spreading operations among S-boxes achieves Shannon's concept of diffusion. A primary key is used to generate all subkeys implemented in each SP layer according to an algorithm referred to as a key schedule. Key schedules will not be discussed in this thesis. In each SP layer, a subkey is either XORed with the input bits of that layer, then fed into the S-boxes, or used to select different mappings for the S-boxes in that layer. Two layers of an SPN based on 4-bit S-boxes are shown in Figure 2.1.

2.1.2 DES-like Block Ciphers

Feistel, et al [12] proposed another cipher structure which was adopted by the National Bureau of Standards (NBS) as the network architecture of DES [24]. Such a structure, illustrated in Figure 2.2, is known as the DES-like cipher architecture.

Unlike SPNs, a DES-like cipher divides an N-bit plaintext into two $N/2$-bit halves referred to as the left and right halves. For each round, the right half is input into a round function F whose output is bit-wise XORed with the left half, then the two

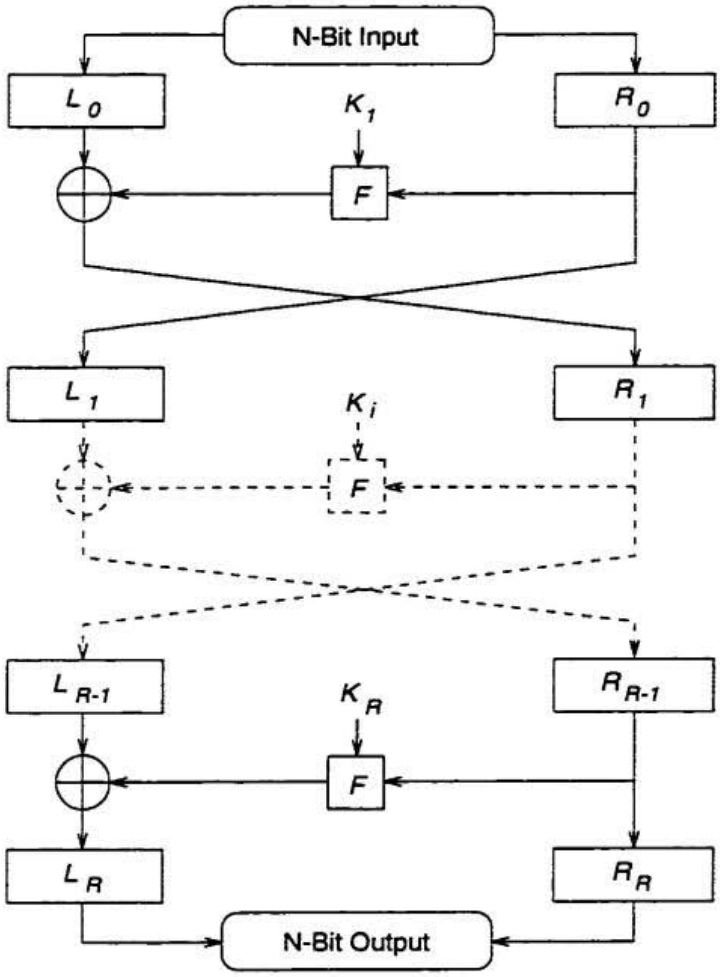

Figure 2.2: The Structure of DES-like Ciphers

halves are swapped. After R rounds, the two halves are concatenated to form the N-bit ciphertext. The cipher can be viewed as the following iterated operation:

$$\begin{aligned} L_i &= R_{i-1} \\ R_i &= L_{i-1} \oplus F(R_{i-1}, K_i), \end{aligned} \tag{2.1}$$

for $1 \leq i \leq R-1$, and

$$R_R = R_{R-1}$$

$$L_R = L_R \oplus F(R_R, K_R), \tag{2.2}$$

where $\oplus$ represents bit-wise XOR.

Encryption and decryption are achieved by the same algorithm. However, for decryption the subkeys are used in reverse order. A key schedule is designed to generate subkey K_i for every round from a cipher key.

The plaintext randomization is performed by the round function, which is the crucial part in DES-like ciphers. The main difference among many DES-like ciphers is the structure of the round function.

DES is a 64-bit block cipher which has 16 rounds and a 56-bit key size. The round function first expands its 32-bit input to a 48-bit block using a bit-selection table. The expanded block is then XORed with a 48-bit subkey generated by the key schedule algorithm. The XORed 48-bit block is finally fed into eight 6×4 S-boxes whose output passes through a 32-bit permutation to be the 32-bit output of the round function.

2.2 Proposed Private-key Block Ciphers

Many private-key block ciphers have been proposed as potential replacements for DES. These ciphers may be DES-like or not. In this section we introduce several algorithms.

FEAL [32, 23] is exactly a DES-like encryption algorithm. It is a 64-bit block cipher with a 64-bit key which can be easily and efficiently implemented in the 8-bit microprocessor environment. The 8×8 S-boxes in the round function perform byte

rotations and XOR additions. Unfortunately, FEAL with less than 8 rounds was easily broken by differential cryptanalysis [4], and requires at least 32 rounds to be resistant to differential cryptanalysis [26].

LOKI [7] is also a 64-bit block cipher with a 64-bit key similar to DES. The round function employs 12 × 8 S-boxes based on irreducible polynomials, and re-arranges the expansion and permutation tables. The initial version of LOKI was found to be susceptible to differential cryptanalysis [5], and has since been strengthened [6].

IDEA [17] is a 64-bit block cipher with a 128-bit key. It is not exactly a DES-like cipher in nature. The concepts of confusion and diffusion are achieved by mixing three different groups of operations – bit by bit exclusive-OR, addition of integers, and multiplication of integers. It is claimed that the improved version of the cipher is easily implemented in software and hardware, and resistant to differential cryptanalysis.

Khafre [21] is a software-oriented encryption algorithm with 64-bit blocks whose number of rounds is not specified, but should be a multiple of eight. The round function employs an 8 × 32 S-box to perform the plaintext randomization, and a rotation schedule brings every byte of the plaintext to the input of the S-box once every eight rounds. Biham successfully applied differential cryptanalysis to break 16-round Khafre with 1536 encryptions [5].

RC5 [28] is a fast word-oriented, symmetric block cipher suitable for software implementations. It is a parameterized family of encryption algorithms, which has a variable word size, a variable number of rounds, and a variable-length secret key. Different from the concept of SPNs, the novel feature of RC5 is the heavy use of

data-dependent rotations in which one word of intermediate data is cyclically rotated by an amount determined by the low-order bits of another word of intermediate data. It appears to frustrate differential cryptanalysis and linear cryptanalysis [15].

2.3 Cryptographic Properties

In this section, we present several cryptographic properties which are important to design and analyze the S-box and the whole cipher structure.

2.3.1 Nonlinearity

If some of the plaintext bits, ciphertext bits, and key bits have linear relations, the cipher could be easily broken by solving a set of linear equations with a small amount of known plaintext-ciphertext pairs. Since S-boxes are the only nonlinear components in SPNs and DES-like ciphers, the design of highly nonlinear S-boxes becomes crucial to the development of highly secure ciphers.

An m-bit affine boolean function is defined to be a function of the form

$$A(X) = a_0 \oplus a_1 x_1 \oplus \ldots \oplus a_m x_m \tag{2.3}$$

where $X = [x_1, \ldots, x_m]$ represents the m-bit binary input and $a_i \in \{0, 1\}$, $0 \leq i \leq m$. The Hamming distance between two m-bit boolean functions, $f(X)$ and $g(X)$, can be defined to be

$$d(f, g) = \#\{X \in \{0, 1\}^m | f(X) \oplus g(X) = 1\}. \tag{2.4}$$

Then the nonlinearity of an m-bit boolean function f is defined as

$$N(f) = \min_{g \in \mathcal{A}} d(f, g) \tag{2.5}$$

where $\mathcal{A}$ is the set of all m-bit affine boolean functions. Since an $m \times n$ S-box has n output bits, each of which is an m-bit boolean function, the nonlinearity of the S-box S is defined as the minimum nonlinearity over all non-zero linear combinations of output bit boolean functions:

$$\mathcal{N}(S) = \min_{c_i \in \{0,1\},\ not\ all\ c_i = 0} \mathcal{N}(\bigoplus_{i=1}^{n} c_i f_i) \tag{2.6}$$

where f_i is the m-bit boolean function of the i-th output bit of the S-box.

The maximum nonlinearity of an m-bit boolean function for even values of m is given by [20]

$$N_{\max} = 2^{m-1} - 2^{m/2-1}. \tag{2.7}$$

Only a special class of functions referred to as bent functions [29] can achieve maximum nonlinearity. Nyberg [25] proves that an $m \times n$ S-box can be perfectly nonlinear (i.e. $\mathcal{N}(S) = N_{\max}$) if and only if $m \geq 2n$.

2.3.2 Information Theory

The basic concepts of information theory, such as entropy, mutual information, equivocation, and redundancy were first introduced by Shannon [31] to analyze the security of cryptosystems. Let P be the plaintext and C be the ciphertext. If the conditional entropy $H(P|C)$ is equal to the entropy $H(P)$, then the cipher has *perfect secrecy*. Unfortunately, perfect secrecy is generally impractical to achieve.

Forré [13] first presented a set of cryptographic properties of S-boxes based on information theory. Dawson and Tavares [10] extended Forré's ideas to define an expanded set of design criteria for cryptographically strong S-boxes. Sivabalan, et al

[33] developed Dawson and Tavares's information leakage from a single bit level to a multiple bit level.

Let two random variables X and Y have possible values $X \in \{x_1, \ldots, x_m\}$ and $Y \in \{y_1, \ldots, y_n\}$ separately. The uncertainty or entropy of Y is defined as

$$H(Y) = -\sum_{i=1}^{n} p(y_i) \log p(y_i) \tag{2.8}$$

where $p(y_i)$ is the probability of $Y = y_i$. The conditional entropy of Y given X is defined as

$$H(Y|X) = -\sum_{i=1}^{m} \sum_{j=1}^{n} p(x_i, y_j) \log p(y_j|x_i) \tag{2.9}$$

where $p(x_i, y_j)$ is the joint probability of $X = x_i$ and $Y = y_j$, and $p(y_j|x_i)$ is the conditional probability of $Y = y_j$ given $X = x_i$. The base of the logarithm is arbitrary and amounts to a constant multiplicative factor.

The static input-output information leakage of an S-box is the mutual information, $I(Y, X)$, and is defined as

$$I(Y, X) = H(Y) - H(Y|X) \tag{2.10}$$

where X is any subset of the input bits, and Y is any subset of output bits of an S-box. The dynamic input-output information leakage is the mutual information, $I(\Delta Y, \Delta X)$, and is defined as

$$I(\Delta Y, \Delta X) = H(\Delta Y) - H(\Delta Y|\Delta X) \tag{2.11}$$

where ΔX is XOR changes in a subset of input bits, and ΔY is XOR changes in a subset of output bits of an S-box.

An ideal S-box should have both static and dynamic input-output information leakage equal to zero. However, due to the deterministic nature of an S-box whose input-output mappings are known, an ideal S-box can not be obtained. The cryptographically strong S-box must have as small an information leakage as possible. This is often an objective of designing a private-key block cipher.

The static and dynamic information leakage reflect the vulnerability of an S-box to correlation attacks, differential cryptanalysis and linear cryptanalysis. In [37], Zhang, Tavares and Campbell suggested that the information leakage can be used as a fundamental measure of the strength of an S-box and an encryption algorithm comprehensively, instead of all other cryptographic criteria which generally only reflect one aspect of vulnerability separately, such as nonlinearity, higher order Strict Avalanche Criterion (SAC), and correlation immunity.

2.3.3 Other Cryptographic Properties

Other cryptographic properties are avalanche [11, 12], completeness [16] and the SAC [34]. For a cipher, the avalanche criterion is strictly satisfied if, on average, half of the ciphertext bits will change when one plaintext bit changes, and the completeness criterion is satisfied if all output bits depend on all input bits. In [8], Brown and Seberry have found that DES is complete after four to five rounds with a high probability.

Webster and Tavares combined the avalanche and completeness criteria into the SAC. A cipher satisfies SAC if a one bit plaintext change causes each ciphertext bit to change with a probability of 1/2. DES has been found to satisfy SAC after five or six rounds.

2.4 Cryptanalysis

The objective of cryptanalysis is to determine a secret master key used in a cipher. The general classes of cryptanalysis are ciphertext only, known plaintext, and chosen plaintext. A ciphertext only attack has the knowledge of ciphertexts only. A known plaintext attack uses the knowledge of both plaintexts and corresponding ciphertexts. A chosen plaintext attack assumes that a cryptanalyst can choose specific sets of plaintexts, and obtain the corresponding sets of ciphertexts.

In this section, we briefly introduce the three most powerful and widely used cryptanalysis techniques of private-key block ciphers. Differential and linear cryptanalysis will be described in more detail in the next chapter.

2.4.1 Exhaustive Key Search

The method of exhaustive key search is a known plaintext attack. The cryptanalyst first acquires a known plaintext-ciphertext pair encrypted with an unknown master key and encrypts the known plaintext with all possible keys. When a key generates the correct ciphertext, with high probability, it is the correct key. If necessary, several known plaintext-ciphertext pairs can be used to verify its correctness.

Usually ciphers are designed to have a large enough key size to make exhaustive search infeasible. Unfortunately, the 56-bit key size of DES is so small that it has received extensive criticism. In general, if the work load of a cryptanalytic attack is less than the work load for exhaustive search of the key space, the cipher is theoretically broken.

2.4.2 Differential Cryptanalysis

In [3], Biham and Shamir developed differential cryptanalysis, one of the most powerful cryptanalytic methods on iterated product ciphers, such as SPNs and DES-like ciphers. They have published a series of papers attacking DES, FEAL, LOKI, and other proposed ciphers by differential cryptanalysis [3, 4, 5], which have forced the re-design of several proposed ciphers. Most impressively, they have demonstrated a successful cryptanalysis of 16-round DES with 2^{47} chosen plaintexts [3].

Differential cryptanalysis is a chosen plaintext attack which compares the bit-wise XOR value of two plaintexts to the XOR value of the corresponding two ciphertexts. In an S-box, the knowledge of the input XOR of a pair cannot guarantee the knowledge of its output XOR. However, every input XOR of an S-box suggests a probabilistic distribution of the possible output XORs. Given a particular input XOR, it is possible for some output XORs to have a relatively high probability. It is such high probabilities that can be utilized to exploit multiple round XOR structures called differential characteristics with a high probability and derive a portion of the subkey bits applied in an R-round cipher.

In order to make the round function immune to differential cryptanalysis, several methods have been proposed. One is to reduce high probabilities in XOR distributions of S-boxes, which can be achieved by extending the number of output bits of the S-boxes to a reasonable value [2]. Another approach is to replace the XOR operation in the round function which involves the subkey by a modular multiplication to hide the input of the S-boxes [1].

2.4.3 Linear Cryptanalysis

In [19], Matsui introduced a known plaintext attack against DES: linear cryptanalysis. It studies statistical linear relations between bits of plaintexts, ciphertexts and subkeys. By this method, the full 16-round DES cipher is broken with 2^{47} known plaintext-ciphertext pairs.

Truly, there are many statistical linear relations referred to as linear approximations between input and output bits of an S-box. Since all operations in DES, except the S-boxes, are linear, these linear relations can be utilized to construct a linear approximation of the entire algorithm. Matsui presented two algorithms to derive the subkey bits from a linear approximation, based on hypothesis testing. Algorithm 1 can retrieve an equivalent subkey bit expressed as the XOR sum of the subkey bits. Algorithm 2 can more efficiently find a number of the subkey bits at one time.

It is obvious that choosing S-boxes with high nonlinearity is an effective way to make SPNs resistant to linear cryptanalysis. Introducing nonlinear operations in round functions is another way to make ciphers immune to linear cryptanalysis, such as key-dependent rotations [28] and modular additions and subtractions [30].

2.5 Design of CAST components

Adams and Tavares [2] proposed a new DES-like cryptosystem called the CAST encryption algorithm, which seems promising in resistance to differential and linear cryptanalysis, and has high encryption/decryption performance.

In this section, we introduce the CAST design procedure, focussing mainly on the

structure of the S-box and the round function, which are significantly different from DES.

2.5.1 Substitution Box

As we mentioned, the S-box in an SPN or DES-like cipher is very important to security since it is the only source of nonlinearity. The CAST design procedure makes use of S-boxes which have fewer input bits than output bits, such as 8×32 S-boxes. Some research has indicated that this class of S-boxes exhibits good confusion, complete diffusion, good avalanche, highest-order SAC, a flat distribution of output XORs, and high nonlinearity [22]. These important cryptographic properties directly influence the security and efficiency of the entire algorithm. Although confusion, diffusion, and avalanche are somewhat abstract concepts and cannot be proven formally, the CAST cipher constructed by such S-boxes shows good statistical properties after 2-3 rounds while DES needs 5-6 rounds to exhibit similar properties.

2.5.2 Round Function

Round functions of DES and CAST with a 64-bit blocksize are shown in Figure 2.3, where E and P are the expansion and permutation functions of DES. In the round function of CAST, 32-bit data R_{i-1} is input to the round function and XORed with a subkey K_i. The 32-bit result is split into four 8-bit pieces each of which is input to a different 8×32 S-box and the four 32-bit S-box outputs are XORed together to form the 32-bit modified data, $F(R_{i-1}, K_i)$. Although each S-box causes data expansion, the structure of the round function guarantees that there is no data expansion by the

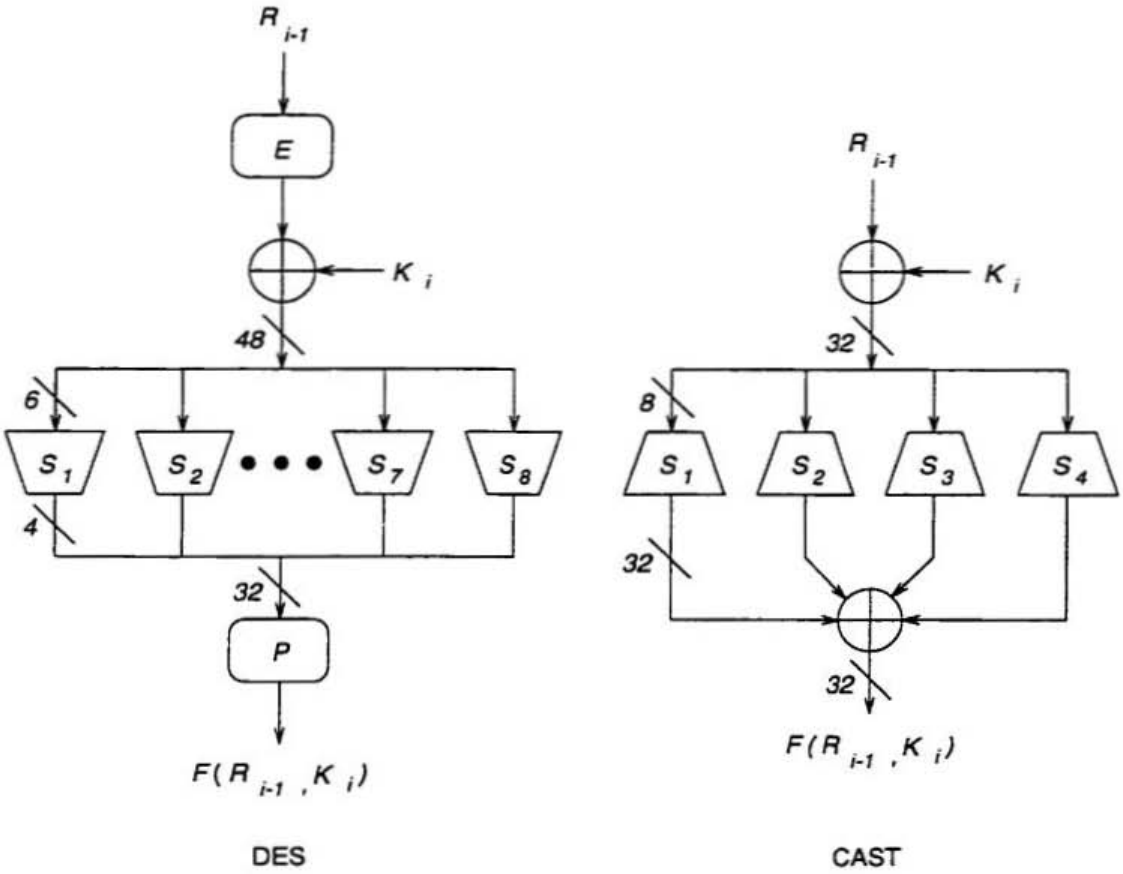

Figure 2.3: The Round Functions of DES and CAST

round function operation.

We have noticed that CAST eliminates the permutation layer in its round function because of using four large 8×32 S-boxes. In the round function of DES, the permutation serves to spread output bit changes from a single small S-box over the block half so that these changes are input to several small S-boxes in the next round. In CAST, each S-box directly affects the entire block half so that any output bit change is guaranteed to affect all S-boxes in the next round without a permutation layer. This modification greatly benefits the software implementation and improves the encryption/decryption performance because the permutation is inefficiently im-

plemented in software. CAST has been shown to be two to three times faster than a typical implementation of DES [1].

2.6 Cryptanalysis of CAST

In this section, we describe three different attacks which have been applied to the CAST encryption algorithm, differential and linear cryptanalysis, and the attack based on non-surjective round functions.

2.6.1 Differential Cryptanalysis of CAST

Lee, et al [18] showed that CAST, using randomly generated S-boxes, is resistant to differential cryptanalysis. In their paper, they developed a method to predict the entry distribution of the XOR table of the round function of CAST. Although each S-box has a highest differential probability of 2^{-7}, a simple screening process can be applied to prevent the occurrence of a 2-round iterative characteristic with a probability of 2^{-7}. They claimed that the best 2-round iterative characteristic has a probability of 2^{-14} for CAST, while the corresponding probability in DES is 1/234. In general, a $(R-2)$-round characteristic can be used to attack an R-round cipher. Using a concatenation of the four best 2-round iterative characteristics, a 10-round CAST has a characteristic with a probability of 2^{-56}, which is better than one for 16-round DES.

2.6.2 Linear Cryptanalysis of CAST

Heys and Tavares [14] investigated the security of CAST with respect to linear cryptanalysis. Assuming that all four 8×32 S-boxes have nonlinearities greater than or

equal to 64, the analysis shows that at least 2^{50} known plaintext-ciphertext pairs are required for linear cryptanalysis to determine only one equivalent key bit of a 12-round CAST, compared to 2^{47} known plaintext-ciphertext pairs required to determine all key bits of 16-round DES. Furthermore, it is a very difficult task for a cryptanalyst to find a linear approximation close to the lower bound.

The result also suggests that at least 99.95% of all randomly generated 8×32 S-boxes have nonlinearities of at least 64. In recent experiments, all randomly generated 8×32 S-boxes have been found to have nonlinearities greater than 72 [36].

2.6.3 Attack Based on Non-surjective Round Functions

Rijmen and Preneel [27] suggested a statistical attack on DES-like ciphers with non-surjective or non-uniform round functions. For a DES-like cipher with R rounds, the following equation holds:

$$\beta_R(L_0, R_0, K) \stackrel{def}{=} \bigoplus_{i=1}^{R/2} F_{2i}(K_{2i} \oplus R_{2i-1}) = R_0 \oplus L_R. \tag{2.12}$$

By rearranging the terms in Equation 2.12, then

$$\beta_{R-2}(L_0, R_0, K) = \bigoplus_{i=1}^{R/2-1} F_{2i}(K_{2i} \oplus R_{2i-1}) = R_0 \oplus L_R \oplus F_R(K_R \oplus R_R). \tag{2.13}$$

If the number of rounds R is small, non-surjective round functions F_{2i} will result in a non-surjective β_{R-2}. A basic attack can be carried on by calculating the right hand side of Equation 2.13 using the known plaintext R_0 and ciphertext L_R for all values K_R. Wrong key candidates will eventually produce a value which is outside the range of β_{R-2}. Even if the number of round R becomes larger and β_{R-2} becomes

surjective, β_{R-2} will not be uniformly distributed. A statistical attack can still be applied to derive the most probable keys.

The round function of CAST is constructed by four 8×32 S-boxes. The output is obtained by XORing the outputs of four S-boxes each of which only has 256 outputs out of 2^{32} all possible values. If the four S-boxes are selected randomly, the expected number of possible outputs are $(1 - e^{-1}) \times 2^{32}$, where e denotes the natural logarithm base, about 63% of all possible values.

The basic attack can be applied to 6-round CAST, requiring a work factor of 1.5×2^{48} operations and 82 known plaintext-ciphertext pairs. However, this attack is not applicable to CAST with more than six rounds since the XOR sum of two CAST round functions is surjective. The statistical attack has not been implemented due to the requirement of a table of size 2^{32}.

2.7 Conclusion

We have introduced two main structures of encryption algorithms used in private-key block ciphers, and several proposed block ciphers. In addition, we have presented several cryptographic properties that are crucial to design and analyze S-boxes and ciphers. Furthermore, we have briefly introduced the two most powerful cryptanalysis methods of private-key block ciphers. Finally, we have outlined the CAST encryption algorithm design procedure and discussed the security of the CAST with respect to three proposed attacks.

Chapter 3

Differential and Linear Cryptanalysis

In this chapter, we describe differential and linear cryptanalysis in more detail. Both attacks have successfully been applied to a variety of SPNs or DES-like ciphers and have forced those ciphers to be redesigned to enhance their security. Differential and linear cryptanalysis appear to be fairly general-purpose attacks and help to qualify the design parameters of private key block ciphers.

3.1 Differential Cryptanalysis

Differential cryptanalysis [3] is a chosen plaintext attack which utilizes highly probable occurrences of output XOR differences of each round function, given particular input XOR differences. In this section we first introduce definitions of XOR tables and characteristics, and then describe the general attack.

3.1.1 XOR Table

A table which shows the distribution of the input XORs and output XORs of all the possible pairs of an S-box is called the XOR table. In the XOR table, each

row corresponds to a particular input XOR value, each column corresponds to a particular output XOR value. For a given $m \times n$ S-box constructed from a mapping $S : \{0,1\}^m \rightarrow \{0,1\}^n$, letting $\Delta X \in \{0,1\}^m$ and $\Delta Y \in \{0,1\}^n$, the XOR table entry of the S-box corresponding to $(\Delta X, \Delta Y)$ is defined as

$$XOR(\Delta X, \Delta Y) = \#\{X \in \{0,1\}^m | S(X) \oplus S(X \oplus \Delta X) = \Delta Y\} \qquad (3.1)$$

where # denotes the cardinality of the set and $\oplus$ represents bit-wise XOR. The XOR table has the following features:

- A zero input XOR always generates a zero output XOR, since the same input value will map to the same output value.
- Entries in an XOR table are always multiples of 2, because two input pairs $(X, X \oplus \Delta X)$ and $(X \oplus \Delta X, X)$ will yield the same output difference ΔY.
- Not all entries are possible. For example, about 20% of all entries are impossible in a 6×4 S-box of DES.

Consider Figure 3.1. Let X and X^* be two input data blocks, Y and Y^* be two corresponding output data blocks, and K be a subkey. Then we have $I = X \oplus K$ and $I^* = X^* \oplus K$. It is obvious that

$$\Delta I = \Delta X \qquad (3.2)$$

where $\Delta X = X \oplus X^*$, $\Delta I = I \oplus I^*$. Equation 3.2 means that the actual input XOR to the S-box, ΔI, does not depend on the subkey. However, the output XOR $\Delta Y = Y \oplus Y^*$ does depend on the subkey because both Y and Y^* are related to the subkey in a nonlinear fashion.

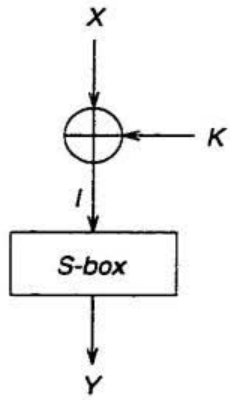

Figure 3.1: Input and Output of an S-box

If a triple $(X, X^*, \Delta Y)$ is known, the XOR table will suggest that there are $XOR(\Delta X, \Delta Y)$ possible input values to the S-box, which can derive a set of possible subkeys by $K = X \oplus I$. Given many such triples, the correct subkey is suggested by all triples.

3.1.2 Characteristic

Let the input and output XOR differences of the i-th round function be ΔX_i and ΔY_i, respectively. Then the probability with which ΔX_i may cause ΔY_i is denoted by p_i. An r-round characteristic, Ω_r, is defined as a sequence of XOR difference pairs $\Omega_r = \{(\Delta X_1, \Delta Y_1), \ldots, (\Delta X_r, \Delta Y_r)\}$. From Figure 3.2, all XOR pairs of the characteristic satisfy the following requirements:

$$\begin{aligned} \Delta X_1 &= \Delta R_0 \\ \Delta X_2 &= \Delta L_0 \oplus \Delta Y_1 \\ &\vdots \quad \vdots \\ \Delta X_i &= \Delta X_{i-2} \oplus \Delta Y_{i-1} \end{aligned}$$

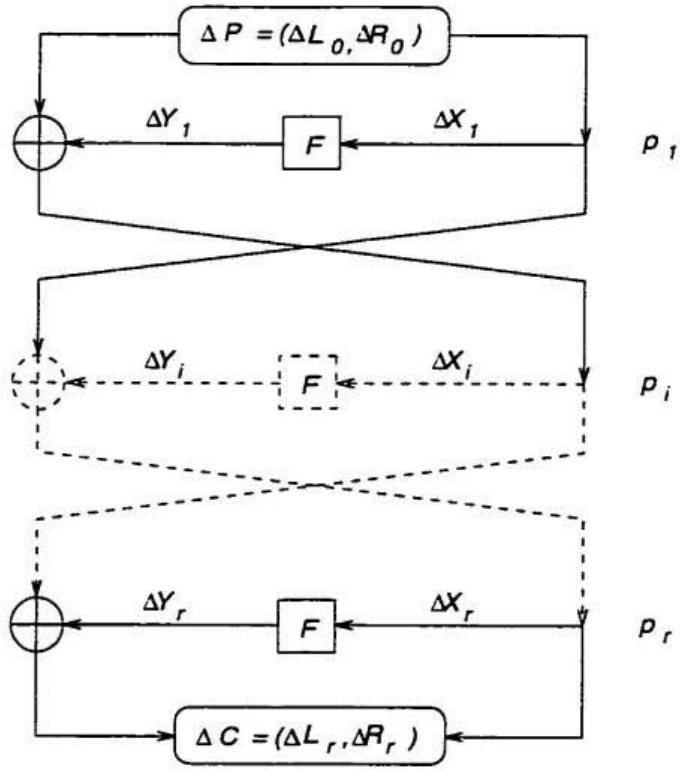

Figure 3.2: r-Round Characteristic Ω_r

$$\begin{aligned} \vdots \quad & \vdots \\ \Delta X_{r-1} &= \Delta L_r \oplus \Delta Y_r \\ \Delta Y_r &= \Delta R_r \end{aligned} \tag{3.3}$$

where the input and output XORs of the characteristic is denoted by $(\Delta L_0, \Delta R_0)$ and $(\Delta L_r, \Delta R_r)$. The probability of the r-round characteristic P_{Ω_r} is given by

$$P_{\Omega_r} = \prod_{i=1}^{r} p_i, \tag{3.4}$$

assuming independence between rounds.

A pair of plaintexts with XOR $\Delta P = (\Delta L_0, \Delta R_0)$ is defined as a right pair with respect to a characteristic Ω_r, if all XOR pairs $(\Delta X_i, \Delta Y_i)$ satisfy Equation 3.3 for

$1 \leq i \leq r$. Otherwise, the pair is defined as a wrong pair. The right pair produces $(\Delta L_r, \Delta R_r)$ from $(\Delta L_0, \Delta R_0)$ with the probability P_{Ω_r}. It is obvious that, for the given input XOR $(\Delta L_0, \Delta R_0)$, a right pair is more likely to occur with fewer number of chosen plaintexts if P_{Ω_r} is high. The objective of differential cryptanalysis is to find a characteristic with a relatively high probability.

The most useful characteristic is one called an iterative characteristic which can be concatenated with itself. The advantage of iterative characteristics is that we can build an r-round characteristic for any large r with a fixed reduction rate of the probability for each additional round, while in non-iterative characteristics the reduction rate of the probability usually increases due to the avalanche effect. The 2-round iterative characteristic is illustrated in Figure 3.3, based on a non-zero XOR input to the round function which may cause a zero XOR output with a relatively high probability.

3.1.3 General Attack

In general, differential cryptanalysis attempts to find the subkey bits used in the last round of DES-like ciphers. This may be achieved if the cryptanalyst is aware of a characteristic with a high probability for the first $(R-1)$ rounds, Ω_{R-1}, and targets the round R S-boxes with non-zero input XORs. In the attack, the cryptanalyst has the knowledge of the exact values of the input pairs to the round function of the last round, R_R and R_R^*, which are the right halves of the ciphertexts. Although the corresponding output XOR value of the round function is not known directly, when a right pair occurs it can be derived by XORing the right half output XOR value of the

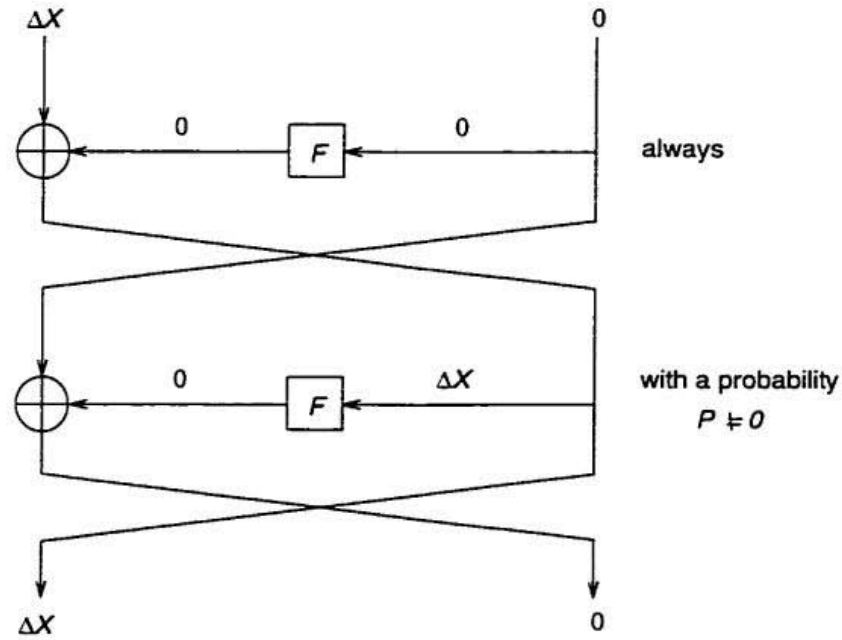

Figure 3.3: 2-Round Iterative Characteristic

$(R-1)$ characteristic, ΔR_{R-1}, with the left half XOR value of the ciphertexts, ΔL_R. Using the method described in Section 3.1.1 and detailed in [3], we can get a set of possible subkey values used in the last round. The correct subkey value should be the one occurring most frequently when a sufficient number of chosen plaintext pairs are considered. Occurrences of the other possible subkey values should be fairly randomly distributed. Some wrong pairs can be directly discarded by checking whether the right half XOR values of the ciphertexts are equal to the expected left half output XOR values of the $R-1$ characteristic, Ω_{R-1}.

3.1.4 Counting Scheme

A counting scheme is used to count the number of occurrences of all possible subkey values used in the last round and finally get the correct subkey values. The following

definition is usable to evaluate a counting scheme based on a characteristic.

Definition 3.1 [3] *The ratio between the number of right pairs and the average count of the incorrect subkeys in a counting scheme is called the signal to noise ratio of the counting scheme and is denoted by SNR.*

The magnitude of *SNR* is determined by the following factors:

- p, the differential probability of a characteristic. Obviously, a high probability characteristic generates several right pairs with a few chosen plaintext XOR pairs.
- k, the number of subkey bits which we simultaneously count on. Counting on a large number of subkey bits simultaneously is helpful to identify the correct key values and needs a small amount of data. However, it demands more memory, which can make the attack impractical.
- γ, the average count among all chosen plaintext XOR pairs. If we can distinguish more wrong pairs by the characteristic, we need fewer counts to get the correct key values. In general, it is difficult to determine the value of γ.

The signal to noise ratio of a counting scheme, *SNR*, is

$$SNR = \frac{2^k \cdot p}{\gamma}. \tag{3.5}$$

Furthermore, we define the number of chosen plaintexts, N_D, required to uniquely identify the correct value of the subkey as

$$N_D = \frac{\eta}{p} \tag{3.6}$$

where η is the number of right pairs required to uniquely determine the correct value of the subkey. When SNR is high, η will be small. When SNR is low, η will become large. Unfortunately, the exact relationship between SNR and η is unknown. However, no matter what value SNR is, η should be greater than one.

3.2 Linear Cryptanalysis

Linear cryptanalysis [19] is a known plaintext attack which extracts key information by finding a statistical linear equation consisting of plaintext, ciphertext, and key terms only. In this section we introduce the basic idea of linear cryptanalysis.

3.2.1 Basic Attack

Usually, the linear approximation has the form of

$$P_{i_1}\oplus,\ldots,\oplus P_{i_a}\oplus C_{j_1}\oplus,\ldots,\oplus C_{j_b} = K_{k_1}\oplus,\ldots,\oplus K_{k_c} \tag{3.7}$$

where $i_1,\ldots,i_a$, $j_1,\ldots,j_b$, $k_1,\ldots,k_c$ denote fixed bit positions of the plaintext P, ciphertext C, and key K, respectively. Let P_L be the probability with which Equation 3.7 holds. If $|P_L - 1/2|$ is large enough and sufficient plaintext-ciphertext pairs are known, it is possible to use a hypothesis test to determine one equivalent key bit which is expressed by the XOR sum of the key bits on the right hand side of Equation 3.7.

Let N_L be the number of given random plaintexts. Matsui's Piling-up Lemma [19] shows that one equivalent key bit of Equation 3.7 can be determined with 97.7% confidence if N_L satisfies

$$N_L \approx |P_L - 1/2|^{-2}. \tag{3.8}$$

3.2.2 Linear Approximation of the S-box

The linear approximations of an S-box are found from the linear approximation table of the S-box. In the linear approximation table, each row corresponds to a subset of input bits of the S-box, each column corresponds to a subset of output bits of the S-box. For a given $m \times n$ S-box constructed from a mapping $S : \{0,1\}^m \rightarrow \{0.1\}^n$, letting $\alpha \in \{0,1\}^m$ and $\beta \in \{0,1\}^n$, the linear approximation table entry of the S-box corresponding to (α, β) is defined as

$$LAT(\alpha, \beta) = \#\{X \in \{0,1\}^m | \alpha X = \beta S(X)\} - 2^{m-1} \tag{3.9}$$

where # denotes the cardinality of the set, and αX is defined as

$$\alpha X = a_1 x_1 \oplus a_2 x_2 \oplus \ldots \oplus a_m x_m \tag{3.10}$$

where a_i and $x_i \in \{0,1\}$, $1 \leq i \leq m$. It is straightforward to notice that

$$LAT(\alpha, \beta) = 2^{m-1} - d(\alpha X, \beta S(X)) \tag{3.11}$$

where

$$d(\alpha X, \beta S(X)) = \#\{X \in \{0,1\}^m | \alpha X \oplus \beta S(X) = 1\}. \tag{3.12}$$

Equation 3.12 is the Hamming distance between an affine boolean function αX and a linear combination of output bit boolean functions of the S-box $\beta S(X)$.

The maximum $|LAT(\alpha, \beta)|$ in the linear approximation table expresses the best linear approximation of the S-box whose probability p is defined as

$$|p - 1/2| = \max |LAT(\alpha, \beta)|/2^m. \tag{3.13}$$

3.2.3 Linear Approximation of the Cipher

A linear approximation of a cipher is derived by XORing a number of linear approximations of S-boxes such that any intermediate terms (i.e. terms are not plaintext, ciphertext, or key terms) are cancelled. Assuming that there are n such linear approximations whose probabilities are $p_1, \ldots, p_n$, then Matsui's Piling-up Lemma shows that the probability of the linear approximation of the cipher, P_L, has

$$|P_L - 1/2| = 2^{n-1} \prod_{i=1}^{n} |p_i - 1/2| \tag{3.14}$$

From Equation 3.8, it is obvious that N_L can be increased by decreasing $|P_L - 1/2|$. Therefore, Equation 3.14 suggests that selecting S-boxes which have $p_i \rightarrow 1/2$ and increasing the number of linear approximations of S-boxes involved in the overall linear approximation increases the cipher's resistance to linear cryptanalysis.

3.3 Conclusion

In this chapter, we have described two of the most powerful cryptanalysis techniques, differential and linear cryptanalysis in more detail. The major concepts of XOR tables and characteristics in differential cryptanalysis, and linear approximations in linear cryptanalysis were discussed. Their successful application to the DES encryption algorithm and many other proposed ciphers suggest that they provide two methods with which judge the security of private-key block ciphers.

Chapter 4

A New General Class of Unbalanced CAST Ciphers

DES [24] is nearing the end of its useful life. After twenty years, it is now theoretically breakable by differential [3] and linear cryptanalysis [19], and practically breakable using special purpose hardware [35]. Many candidates are proposed for the replacement of DES. Unfortunately, there is no obvious one with acceptable speed and security.

In this chapter, we propose a new general class of encryption algorithms referred to as an unbalanced CAST cipher, and describe its design criteria and procedure. The cryptographic analysis results will be given in later chapters.

4.1 Motivation

A successful cipher should have the following features.

- *Security*: This is the main goal of the cipher design. A successful cipher should be resistant to all proposed cryptanalyses, such as differential and linear cryptanalysis, and be a potential survivor in the future when computing capability increases. Furthermore, there should exist a clear mathematical method so that

the cryptographic strength of the cipher can be easily analyzed and evaluated. A complicated encryption algorithm is not guaranteed to have a strong cryptographic strength.

- *Efficiency*: There is usually a trade-off between a higher level of security and a higher encryption/decryption speed. A larger number of rounds usually provides a product cipher, such as SPN, with a higher security, but implies a lower speed. A successful cipher should have a high encryption/decryption speed under a certain level of security, and allow users to explicitly manipulate the trade-off by selecting a variable number of rounds.
- *Parameterization*: With the increasing growth of telecommunications, encryption technology has been more and more adopted by various applications. This gives rise to the problem that a single encryption algorithm may not efficiently fit all applications since hardware and software environments may be totally different. For example, 8-bit, 16-bit, and 32-bit microprocessors have different data processing abilities. Personal computers, cellular phones, and points of sales (POS) terminals have different memories available for an encryption algorithm. Therefore, a successful cipher should be a family of encryption algorithms characterized by a set of parameters.
- *Ease of Implementation*: DES, which was designed in the 1970's, is explicitly a hardware-oriented encryption algorithm. Its extensive use of permutations is inefficient in software, and its 6×4 S-boxes can not be efficiently implemented by modern computers. A new cipher would be most valuable if it is easily

implemented in both software and hardware.

The original CAST cipher [2] appears to be resistant to differential and linear cryptanalysis [18, 14], and possesses a number of desirable cryptographic properties such as avalanche [11, 12] and SAC [34]. It is easily implemented by software and has a good encryption/decryption performance on 32-bit microprocessors. By using four large 8×32 S-boxes, CAST eliminates the need of permutations. However, large S-boxes require more memory to store their lookup tables. This might be unacceptable in some implementations where the memory is extremely restricted, such as smart cards.

Keeping the above objectives in mind, we present a family of ciphers referred to as unbalanced CAST ciphers, which employ the same type of S-box and round function as the original CAST cipher, and have a variable amount of memory requirements dependent on the selection of different parameters.

4.2 Description of the Algorithm

The unbalanced CAST cipher is a product cipher which iterates a round operation R times. The round operation of the general cipher may be conceptualized as in Figure 4.1. Let N be the block size of the cipher. In the i-th round, an N-bit input B_i is first split into three pieces, L_i, H_i, and U_i. L_i is input to the round function F XORed with an l-bit subkey K_i, H_i is XORed with the output of the round function, and U_i bypasses the round function. Finally, L_i, the XOR sum of the output of the round function and H_i, and U_i are processed by a rotation operation and result in an N-bit output, B_{i+1}. The round function has the same structure as the one of

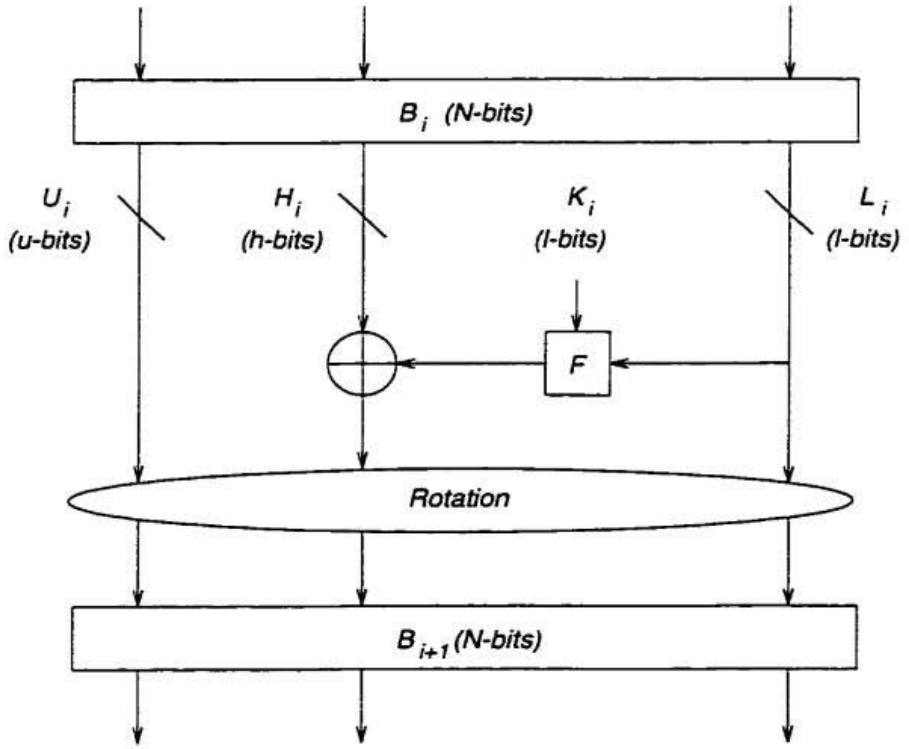

Figure 4.1: i-th Round Operation of an Unbalanced CAST Cipher

the original CAST cipher shown in Figure 2.3. Letting M be the number of $m \times n$ S-boxes used in the round function with $m < n$, the following relationships hold

$$\begin{aligned} l &= M \times m \\ h &= n \\ u &= N - l - h \end{aligned} \tag{4.1}$$

where l is the number of bits of L_i, h is the number of bits of H_i, and u is the number of bits of U_i. Also $l \leq h$.

In general, the round operations of the unbalanced CAST ciphers can be characterized entirely by the parameters N, M, m, n and the rotation operation. For example, the original CAST cipher can be characterized by $N = 64$, $M = 4$, $m = 8$,

$n = 32$, and a 32-bit rotation in the form of swapping the two half blocks. Khafre can be characterized by $N = 64$, $M = 1$, $m = 8$, $n = 32$, and a rotation of eight and sixteen bits specified according to the round number. We refer to the ciphers as unbalanced since l is not necessarily equal to $N/2$. A balanced CAST cipher, such as the original CAST cipher, has $l = n = N/2$ and $u = 0$.

If round functions from two ciphers, Cipher 1 and Cipher 2, are constructed by the same type of $m \times n$ S-boxes and the following equation holds:

$$M_1 \cdot R_1 = M_2 \cdot R_2 \tag{4.2}$$

where the subscript is used to indicate the cipher number, then the two ciphers may be considered to be roughly equivalent in efficiency if the S-box table lookup is considered the dominant operation (i.e. assuming data rotation in CPU registers may be ignored). For example, an 8-round original CAST cipher with four 8×32 S-boxes may be considered roughly equivalent in efficiency to a 32-round unbalanced CAST cipher with one 8×32 S-box in efficiency. However, Equation 4.2 does not imply that two ciphers may have an equivalent level of security. This issue will be addressed in later chapters.

For current computer technology, the blocksize 64 is quite convenient. Therefore, we choose 64 as the default value of N in the remainder of the thesis.

4.3 Design Decisions

Based on the above structure, we now give some design decisions which mainly focus on S-box, rotation operation, and round function.

4.3.1 S-box

An $m \times n$ S-box is a $2^m \times n$ lookup table which maps m input bits to n output bits. In SPNs and most DES-like ciphers, S-boxes are critically important to security since they are the main components of nonlinearity in the algorithm. Since the size of the lookup table will exponentially increase with the increase of the value m, the value m should be chosen to be small, say less than 10. Conversely, the value n can be chosen to be large since the size of the lookup table increases linearly with n.

The original version of CAST uses S-boxes based on bent boolean functions [2]. To simplify the problem and benefit the statistical analysis, we consider randomly generated S-boxes. Research has shown that there is no large cryptographic difference between these two kinds of S-boxes [22]. To construct an $m \times n$ S-box one bit at a time, a random number generator is used whose outputs have either one or zero with a probability of 1/2. The randomly generated 8×32 S-box has following cryptographic properties.

- *Good Avalanche*: Approximately half the output bits change when one input bit changes [11, 12].
- *Bit Independence Criterion*: Any two output bits change independently when any single input bit changes [34].
- *Equiprobable XOR Distribution*: In the XOR table, with a very high probability, all entries are either 0 or 2.
- *High Nonlinearity*: The S-box has a nonlinearity at least 64 with a high probability [14].

Usually, the values m and n are selected as an integer multiple of eight for ease of implementation in modern computers. However, not all S-boxes which have fewer input bits than output bits are suitable to use in the unbalanced CAST ciphers. In later chapters, we will give the analysis results.

4.3.2 Round Function

Differential and linear cryptanalysis work on the principle of finding characteristics and linear approximations with high probabilities on a single round, then cascading a sufficient number of characteristics and linear approximations of different single rounds in useful ways to attack the whole cipher. It can be seen that adding sufficient rounds to a DES-like cipher makes the cipher computationally resistant to these attacks. The disadvantage of this approach is that the encryption/decryption speed of the cipher is reduced. An alternate approach is to decrease the probabilities of characteristics and linear approximations of an individual round by improving the cryptographic properties of S-boxes. This makes the cipher potentially resistant to differential and linear cryptanalysis without losing the efficiency of the algorithm.

The difference of the round functions of DES and the original CAST cipher has been illustrated in Figure 2.3. In general, DES is a hardware-oriented encryption algorithm, while CAST is a software-oriented encryption algorithm. However, a relatively efficient software implementation of DES can eliminate the permutations by regarding all 6×4 S-boxes as 6×32 S-boxes. Then, the outputs of eight 6×32 S-boxes are XORed together in the same way as CAST, and the permutations need not be applied. In this way, DES and CAST can be viewed as having the round functions

similar in structure. The 8×32 S-box of CAST has all output bits influenced by all input bits, whereas the 6×32 S-box of DES has twenty-eight output bits fixed at 0 and only four output bits influenced by all input bits since DES really just has the 6×4 S-box. This feature causes each output bit of the round function of CAST to be influenced by all four 8×32 S-boxes, while the corresponding output bit of DES is influenced by only one of eight 6×32 S-boxes. The simple comparison tells us that the round function of CAST should be stronger than the one of DES.

The unbalanced CAST ciphers make use of the above proposed S-boxes and the same round function structure as CAST, but might have fewer S-boxes in the round function than CAST to reduce the memory requirement. In exchange, more rounds are required to keep the cipher at the same level of security.

4.3.3 Rotation Operation

In the original CAST cipher, a half block passes through the round function in each round. Swapping the two half blocks in each round makes each plaintext bit pass through the round function once after two rounds. However, because the number of input bits of the round function in an unbalanced CAST cipher can be less than the number of output bits, the swapping operation becomes not enough to efficiently achieve completeness.

The primary purpose of the rotation operation is to bring all plaintext bits to the input position of the round function in as few rounds as possible so that each plaintext bit can influence every ciphertext bit in as few rounds as possible. In general, we assume that N is divisible by l. Then it takes N/l rounds to restore the

plaintext bits to their original position and $(N/l+1)$ rounds for $h \geq 32$ to achieve the completeness property since the last l bits of the plaintext will past through the round function in the (N/l)-th round and get influenced by themselves in the $(N/l+1)$-th round. However, if $h < 32$, it will take a few more rounds to achieve completeness since the last l bits of the plaintext will not be immediately influenced by themselves in the $(N/l+1)$-th round. For example, it takes thirteen rounds for an unbalanced CAST cipher with one 8×16 S-box, or seven rounds for an unbalanced CAST cipher with two 8×16 S-boxes to achieve the completeness property.

Based on Figure 4.1, we propose a rotation operation as well as the round operation, which is effective in ensuring that ciphertext bits are influenced by plaintext bits as quickly as possible. It is described as the following:

1. B_i is divided into two halves, the right half and the left half.
2. L_i, taken from the l least significant bits of the right half, is input into the F round function whose output is XORed with H_i, which is the h least significant bits of the left half.
3. The right half is right cyclically rotated by l bits.
4. Two halves are swapped to form B_{i+1}.

Note that this operation is similar to, but slightly inconsistent with the convenient conceptualization of Figure 4.1. The swapping of two half blocks is still necessary in an unbalanced CAST cipher because the H_i XORed with the output of the round function can be immediately brought to the input position of the round function at the next round, which has all ciphertext bits influenced faster by all plaintext bits.

The rotation operation of the 8-round unbalanced CAST cipher with $M = 1$, $m = 8$, and $n = 32$ is illustrated in Table 4.1, where a letter represents an 8-bit data block.

Rnd	Left Half				Right Half			
P_L	A	B	C	D	E	F	G	H
1	A	B	C	D	E	F	G	H
2	H	E	F	G	A	B	C	D
3	D	A	B	C	H	E	F	G
4	G	H	E	F	D	A	B	C
5	C	D	A	B	G	H	E	F
6	F	G	H	E	C	D	A	B
7	B	C	D	A	F	G	H	E
8	E	F	G	H	B	C	D	A
C_L	A	B	C	D	E	F	G	H

Table 4.1: Rotation Operation for an Unbalanced CAST Cipher with $M = 1$, $m = 8$, and $n = 32$

The data to the right side of the double vertical lines is input to the round function whose four byte output data is XORed with the left half block. After every eight rounds the data bits are restored to their original position. It takes nine rounds for the cipher to be complete.

There are two special cases in which the swapping is not necessary, the cipher with $M = 2$, $m = 8$, and $n = 48$, and the cipher with $M = 1$, $m = 8$, and $n = 56$. Both cases have $N = l + h$, where $l < 32$ and $h > 32$. In the first case, $l = 16$ and it takes four rounds before the cipher is complete; in the second case, $l = 8$ and it takes eight rounds before the cipher is complete. The large number of output bits of the S-boxes will make such ciphers suitable for 64-bit processor implementations.

The above rotation operation is a generalized one. A more efficient rotation operation may be used in a specific implementation, depending on the structure of the microprocessor and its registers. For example, Khafre [21], which targets 32-bit micro-

processors with 16-bit registers, uses a combination of eight and sixteen bit rotations to increase its operation speed. However, the main design principle remains same.

4.4 Conclusion

We have presented a new class of encryption algorithms referred to as unbalanced CAST ciphers, which employ the same type of S-box and round function as the original CAST cipher. By choosing different parameters, the cipher can have a variable amount of memory requirements, which is preferred in some implementations. We have also described the general operation of the cipher, and discussed the design decisions of the S-box, the round function, and the rotation operation in detail.

Chapter 5

Differential Cryptanalysis of Unbalanced CAST Ciphers

In this chapter, we examine the resistance of unbalanced CAST ciphers to differential cryptanalysis [3]. An analysis of the distribution of entries in the XOR table of the round function is first developed. Then two methods of cryptanalysis are given, based on Biham's attack [5] and on iterative characteristics. Since it is impossible to analyze all parameters, we only select a set of typical values as a basis for demonstrating the analysis technique and investigating the effect of parameter changes.

As indicated in Section 4.3.1, the 8×32 S-box has some good cryptographic properties, and is utilized by the original balanced CAST cipher ($M = 4$, $m = 8$, and $n = 32$) which has been analyzed in [18, 14]. In this chapter, we will mainly analyze unbalanced CAST ciphers with $m = 8$, $n = 32$ and $M = 1$ or 2. Both ciphers require only 1/4 and 1/2 memory of the original CAST cipher, respectively. For further reduction of the memory requirement, ciphers with parameters of $m = 8$ and $n = 16$ with $M = 1$ or 2, and parameters of $m = 4$ and $n = 32$ with $M = 1$, 2, 4 or 8 will also be investigated. We will also examine the cipher with $M = 1$, $m = 8$, and $n = 56$

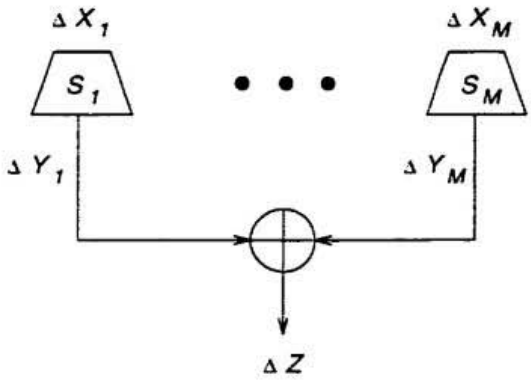

Figure 5.1: XOR Flow in the Round Function of Unbalanced CAST Ciphers

since it is suitable for 64-bit processor implementations.

5.1 Distribution of Entries in the XOR Table

Lee, et al [18] have already given the analysis of the distribution of entries in the XOR table of the round function of the original CAST cipher. The method can be directly used in our analysis.

In order to generalize the analysis, the round function of an unbalanced CAST cipher which is combined by M $m \times n$ S-boxes is shown in Figure 5.1. If we denote the inputs of the S-boxes as $X_1, \ldots, X_M$ and the corresponding outputs of the S-boxes as $Y_1, \ldots, Y_M$, then the output of the round function, Z, is given by

$$Z = \bigoplus_{i=1}^{M} Y_i. \tag{5.1}$$

For the complete round function, given input XOR value $\Delta X = [\Delta X_1, \ldots, \Delta X_M]$, the output XOR value ΔZ is given by

$$\Delta Z = \bigoplus_{i=1}^{M} \Delta Y_i \tag{5.2}$$

where ΔX_i is the bit-wise XOR of two input values, X_i and X_i^*, to the S-box S_i, i.e., $\Delta X_i = X_i \oplus X_i^*$. Also, $\Delta Y_i = Y_i \oplus Y_i^*$ and $\Delta Z = Z \oplus Z^*$.

We define $NZ(\Delta X)$ as the number of S-boxes that have non-zero input XOR values. Then, if $NZ(\Delta X) = M$, all the entries in the XOR table corresponding to that ΔX will be multiples of 2^M. This is because all the following $2^M/2$ input pairs have the same input XOR value ΔX and the same output XOR value ΔZ:

$$\begin{array}{lcl}
\{X_1, X_2, \ldots, X_{M-1}, X_M\} & \leftrightarrow & \{X_1^*, X_2^*, \ldots, X_{M-1}^*, X_M^*\} \\
\{X_1, X_2, \ldots, X_{M-1}, X_M^*\} & \leftrightarrow & \{X_1^*, X_2^*, \ldots, X_{M-1}^*, X_M\} \\
\{X_1, X_2, \ldots, X_{M-1}^*, X_M\} & \leftrightarrow & \{X_1^*, X_2^*, \ldots, X_{M-1}, X_M^*\} \\
\{X_1, X_2, \ldots, X_{M-1}^*, X_M^*\} & \leftrightarrow & \{X_1^*, X_2^*, \ldots, X_{M-1}, X_M\} \\
\vdots & \vdots & \vdots \\
\{X_1, X_2^*, \ldots, X_{M-1}^*, X_M^*\} & \leftrightarrow & \{X_1^*, X_2, \ldots, X_{M-1}, X_M\}
\end{array}$$

Since each of above pairs contributes an entry of 2 in the XOR table corresponding to the row ΔX and the column ΔZ, the $2^M/2$ pairs will result in an entry of 2^M in the XOR table. In a similar way, if $NZ(\Delta X) = i$, all the entries in the XOR table corresponding to that ΔX will be multiples of $2^{m(M-i)} \cdot 2^i$. This is the result of the fact that there are $(M - i)$ $m \times n$ S-boxes that have zero input XOR values and each of the S-boxes has 2^m input pairs that have a zero input XOR value and a zero output XOR value.

For any $m \times n$ S-box, S_i, any input XOR value ΔX_i has $2^m/2 = 2^{m-1}$ unordered pairs of $(X_i, X_i \oplus \Delta X_i)$. If each of these input pairs gives rise to a distinct output pair, then the output XOR ΔY_i can take at most 2^{m-1} distinct values. For the round function, suppose that M $m \times n$ S-boxes are randomly generated and mutually independent, then the output XOR ΔZ can be any one of the 2^n possible values

with equal probability. When $NZ(\Delta X) = M$, the distribution of the output XORs for a given input XOR is equivalent to tossing $(2^{m-1})^M$ balls randomly into 2^n bins with each ball having a weight of 2^M. When $NZ(\Delta X) = i$, the distribution of the output XORs for a given input XOR is equivalent to tossing $(2^{m-1})^i$ balls randomly into 2^n bins with each ball having a weight of $2^{m(M-i)} \cdot 2^i$. We wish to determine the distribution of the balls in the bins in order to examine the distribution of entries in the XOR table.

Let B_k be a random variable representing the number of bins having k balls when a balls are tossed randomly into b bins. It has been shown that if a and b increase in a way such that $b\frac{e^{-\frac{a}{b}}}{k!}(\frac{a}{b})^k$ remains bounded, we can get the expected number of bins with k balls from

$$E(B_k) \approx b\frac{e^{-\frac{a}{b}}}{k!}(\frac{a}{b})^k \tag{5.3}$$

where e is the natural logarithm base.

Consider the round function with four 8×32 S-boxes. If $NZ(\Delta X) = 4$, then $a = 2^{28}$, $b = 2^{32}$, and $E(B_k)$ will be the expected number of ΔZ values each of which has XOR entry of $2^4 \cdot k$ for a particular ΔX. Similarly, if $NZ(\Delta X) = 3$, then $a = 2^{21}$, $b = 2^{32}$, and $E(B_k)$ will be the expected number of ΔZ values each of which has XOR entry of $2^{11} \cdot k$ for a particular ΔX, and so on. The expected distribution of entry values in the XOR table is displayed in Table 5.1.

For the round functions with one or two 8×32 S-boxes (i.e., for ciphers with $M = 1$ or $M = 2$), their distributions of entry values in the XOR table can also be immediately derived from the Table 5.1. For example, the distribution table of the

$NZ(\Delta X)$	k	Entry Value	Number of Entries
4	1	$2^4 \cdot 1$	$2^{27.91}$
4	2	$2^4 \cdot 2$	$2^{22.91}$
4	3	$2^4 \cdot 3$	$2^{17.32}$
4	4	$2^4 \cdot 4$	$2^{11.32}$
4	5	$2^4 \cdot 5$	$2^{5.00}$
4	6	$2^4 \cdot 6$	$2^{-1.58}$
4	7	$2^4 \cdot 7$	$2^{-8.39}$
⋮	⋮	⋮	⋮
3	1	$2^{11} \cdot 1$	$2^{20.99}$
3	2	$2^{11} \cdot 2$	$2^{8.99}$
3	3	$2^{11} \cdot 3$	$2^{-3.59}$
⋮	⋮	⋮	⋮
2	1	$2^{18} \cdot 1$	$2^{14.00}$
2	2	$2^{18} \cdot 2$	2^{-5}
⋮	⋮	⋮	⋮
1	1	$2^{25} \cdot 1$	$2^{7.00}$
1	2	$2^{25} \cdot 2$	2^{-19}
⋮	⋮	⋮	⋮

Table 5.1: Distribution of Entry Values for a Particular ΔX in the XOR Table for $M = 4$, $m = 8$, and $n = 32$

round function with two 8×32 S-boxes begins from $NZ(\Delta X) = 2$. All the numbers in Table 5.1 can be directly quoted except that the entry values should be divided by $2^{m(M-i)} = 2^{16}$. It can be seen that for $M = 1$ and 2, the expected number of bins which have the number of balls greater than 1 for a particular ΔX is less than or equal to 2^{-5}, which is so small that all the bins can be considered to have either 0 or 1 ball. In other words, for $M = 1$ and 2, the maximum likelihood entry values for a particular ΔX are 2 and 4, respectively.

The expected distribution of entry values for a particular ΔX in the XOR table of the round function with two 8×16 S-boxes is displayed in Table 5.2. When

$NZ(\Delta X)$	k	Entry Value	Number of Entries
2	1	$2^2 \cdot 1$	$2^{13.64}$
2	2	$2^2 \cdot 2$	$2^{10.64}$
2	3	$2^2 \cdot 3$	$2^{7.05}$
2	4	$2^2 \cdot 4$	$2^{3.05}$
2	5	$2^2 \cdot 5$	$2^{-1.27}$
⋮	⋮	⋮	⋮
1	1	$2^9 \cdot 1$	$2^{7.00}$
1	2	$2^9 \cdot 2$	2^{-3}
⋮	⋮	⋮	⋮

Table 5.2: Distribution of Entry Values for a Particular ΔX in the XOR Table for $M = 2$, $m = 8$, and $n = 16$

$NZ(\Delta X) = 2$, it is very likely for a particular ΔX that there is a bin having four balls since the expected values are greater than 1. Therefore, the probability that there exist entry values of 4, 8, 12, and 16, is high in the XOR table of the round function.

The expected distribution of entry values for a particular ΔX in the XOR table of the round function with eight 4×32 S-boxes is displayed in Table 5.3. For the round functions with $M = 1$, 2, and 4, it is very unlikely for a particular ΔX that there is a bin having k balls, where $k > 1$. In other words, the maximum entry values for a particular ΔX are 2, 4, and 16, respectively.

In the following section we will utilize the distribution of entry values in the XOR table to derive the likelihood of occurrence of iterative characteristics.

5.2 Iterative Characteristics

The concept of an iterative characteristic has been described in Section 3.1.2. The 2-round iterative characteristic, which is shown in Figure 3.3, is the basic structure

$NZ(\Delta X)$	k	Entry Value	Number of Entries
8	1	$2^8 \cdot 1$	$2^{23.99}$
8	2	$2^8 \cdot 2$	$2^{14.99}$
8	3	$2^8 \cdot 3$	$2^{5.41}$
8	4	$2^8 \cdot 4$	$2^{-4.59}$
8	5	$2^8 \cdot 5$	$2^{-14.91}$
⋮	⋮	⋮	⋮
7	1	$2^{11} \cdot 1$	$2^{21.00}$
7	2	$2^{11} \cdot 2$	$2^{9.00}$
7	3	$2^{11} \cdot 3$	$2^{-3.59}$
7	4	$2^{11} \cdot 4$	$2^{-16.59}$
⋮	⋮	⋮	⋮
6	1	$2^{14} \cdot 1$	$2^{18.00}$
6	2	$2^{14} \cdot 2$	$2^{3.00}$
6	3	$2^{14} \cdot 3$	$2^{-12.59}$
⋮	⋮	⋮	⋮
5	1	$2^{17} \cdot 1$	$2^{15.00}$
5	2	$2^{17} \cdot 2$	$2^{-3.00}$
5	3	$2^{17} \cdot 3$	$2^{-21.58}$
⋮	⋮	⋮	⋮
4	1	$2^{20} \cdot 1$	$2^{12.00}$
4	2	$2^{20} \cdot 2$	$2^{-9.00}$
⋮	⋮	⋮	⋮
3	1	$2^{23} \cdot 1$	$2^{9.00}$
3	2	$2^{23} \cdot 2$	$2^{-15.00}$
⋮	⋮	⋮	⋮
2	1	$2^{26} \cdot 1$	$2^{6.00}$
2	2	$2^{26} \cdot 2$	$2^{-21.00}$
⋮	⋮	⋮	⋮
1	1	$2^{29} \cdot 1$	$2^{3.00}$
1	2	$2^{29} \cdot 2$	$2^{-27.00}$
⋮	⋮	⋮	⋮

Table 5.3: Distribution of Entry Values for a Particular ΔX in the XOR Table for $M = 8$, $m = 4$, and $n = 32$

used for differential cryptanalysis of the full 16-round DES. Lee, Heys and Tavares [18] have found that the best 2-round iterative characteristic of the original balanced CAST cipher (i.e. $M = 4$, $m = 8$, and $n = 32$), has the probability of 2^{-14}. In this section, we statistically analyze the likelihood of occurrence of iterative characteristics for unbalanced CAST ciphers.

5.2.1 Likelihood of Occurrence of Iterative Characteristics

If there is a zero output XOR caused by a non-zero input XOR with a differential probability for a round function, there exists an N/l-round iterative characteristic for that unbalanced CAST cipher. Table 5.4 gives an example of a possible 8-round iterative characteristic for an unbalanced CAST cipher with $M = 1$, $m = 8$, and $n = 32$. The letter A represents a non-zero XOR value and the plaintext XOR

*R*nd	Left Half				Right Half				Output XOR				Probability
ΔP	A	0	0	0	0	0	0	0					
1	A	0	0	0	0	0	0	0	0	0	0	0	
2	0	0	0	0	A	0	0	0	0	0	0	0	
3	0	A	0	0	0	0	0	0	0	0	0	0	
4	0	0	0	0	0	A	0	0	0	0	0	0	
5	0	0	A	0	0	0	0	0	0	0	0	0	
6	0	0	0	0	0	0	A	0	0	0	0	0	
7	0	0	0	A	0	0	0	0	0	0	0	0	
8	0	0	0	0	0	0	0	A	0	0	0	0	p
ΔC	A	0	0	0	0	0	0	0					

Table 5.4: 8-Round Iterative Characteristic for an Unbalanced CAST Cipher with $M = 1$, $m = 8$, and $n = 32$

difference is given by $A0000000$. The XOR values to the right side of the double vertical lines represent input XORs to the round function, while the output XOR column represents the output XORs of the round function. The p is a probability

with which the input XOR value of A may cause the output XOR value of 0000 in the XOR table of the round function.

Note that the characteristic is based on the rotation operation given in Section 4.3.3. The only non-zero XOR byte, A, does not appear as an input to the round function until round 8. Many other iterative characteristics exist and will be described in the next section. Our objective is to find an iterative characteristic with reasonable likelihood of occurring given an appropriate input XOR.

In Section 5.1, we predict the distribution of entry values for a particular ΔX in the XOR table with $NZ(\Delta X) = i$, where $i = 1, \cdots, M$. Now we predict the expected number of entries in the XOR table which can be used to configure an iterative characteristic. Denote $N_E(\Delta X, \Delta Z)$ as an entry in the XOR table of the round function corresponding to an input XOR, ΔX, and an output XOR, ΔZ. Then, $N_E(\Delta X, 0)$ represents an entry corresponding to a particular non-zero input XOR value, ΔX, with $NZ(\Delta X) = i$ and a zero output XOR value, $\Delta Z = 0$. We also denote v as an entry value, which has

$$v = 2^{m(M-i)} \cdot 2^i \cdot k \tag{5.4}$$

where $k = 0, 1, 2, \cdots$. Letting λ_v be the probability with which $N_E(\Delta X, 0) = v$ with $NZ(\Delta X) = i$, then

$$\lambda_v = \frac{E(B_k)}{2^n}. \tag{5.5}$$

If we assume that the occurrence of ΔZs for different ΔXs are independent, since there are totally $\begin{pmatrix} M \\ i \end{pmatrix} \cdot (2^m - 1)^i$ non-zero ΔXs with $NZ(\Delta X) = i$, the expected number of entries in the XOR table which have $N_E(\Delta X, 0) = v$ with $NZ(\Delta X) = i$,

$NZ(\Delta X)$	v	λ_v	μ_v	p
4	$2^4 \cdot 1$	$2^{-4.09}$	2.48×10^8	2^{-28}
4	$2^4 \cdot 2$	$2^{-9.09}$	7.76×10^6	2^{-27}
4	$2^4 \cdot 3$	$2^{-14.68}$	1.61×10^5	$2^{-26.4}$
4	$2^4 \cdot 4$	$2^{-20.68}$	2.52×10^3	2^{-26}
4	$2^4 \cdot 5$	$2^{-27.00}$	3.16×10^1	$2^{-25.7}$
4	$2^4 \cdot 6$	$2^{-33.58}$	3.29×10^{-1}	$2^{-25.4}$
4	$2^4 \cdot 7$	$2^{-40.39}$	2.94×10^{-3}	$2^{-25.2}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
3	$2^{11} \cdot 1$	$2^{-11.00}$	3.24×10^4	2^{-21}
3	$2^{11} \cdot 2$	$2^{-23.00}$	7.90	2^{-20}
3	$2^{11} \cdot 3$	$2^{-35.59}$	1.29×10^{-3}	$2^{-19.4}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
2	$2^{18} \cdot 1$	2^{-18}	1.49	2^{-14}
2	$2^{18} \cdot 2$	2^{-37}	2.84×10^{-6}	2^{-13}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
1	$2^{25} \cdot 1$	2^{-25}	3.04×10^{-5}	2^{-7}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

Table 5.5: Likelihood of Occurrence of Iterative Characteristics for an Unbalanced CAST Cipher with $M = 4$, $m = 8$, and $n = 32$

μ_v, is given by

$$\mu_v = \binom{M}{i} \cdot (2^m - 1)^i \cdot \lambda_v. \tag{5.6}$$

If there is an $N_E(\Delta X, 0) = v$ with $NZ(\Delta X) = i$, where $v \neq 0$ or $k \neq 0$, then there exists an iterative characteristic with a differential probability of p given by

$$p = \frac{v}{2^{Mm}} = k \cdot 2^{-(m-1)i}. \tag{5.7}$$

Table 5.5 gives the likelihood of occurrence of iterative characteristics for a balanced CAST cipher with $M = 4$, $m = 8$, and $n = 32$. It can be seen that it is highly unlikely ($\mu_v = 3.04 \times 10^{-5}$ for $v = 2^{25}$) to construct an iterative characteristic with a differential probability of 2^{-7} for which only one of four S-boxes has an entry corre-

sponding to a non-zero input XOR and a zero output XOR in its XOR table. In fact, it is not difficult to apply a screening process to discard any such S-boxes. However, it is very likely ($\mu_v = 1.49$ for $v = 2^{18}$) that it is possible to construct an iterative characteristic with a differential probability of 2^{-14}. It seems impossible to discard such S-boxes to prevent this iterative characteristic from occurring by a screening process since the likelihood of occurrence is high. Furthermore, for any two of four S-boxes in the round function, $2^{16} \cdot 2^{14} = 2^{30}$ output XOR pairs need to be checked. Totally $\binom{4}{2} \cdot 2^{30} \approx 2^{32.6}$ output XOR pairs need to be checked for the round function. The work load of the screening process is heavy. Therefore, it is assumed that the best iterative characteristic for a balanced CAST cipher with $M = 4$, $m = 8$, and $n = 32$ is a 2-round iterative characteristic with differential probability of 2^{-14}. One of such iterative characteristics is illustrated in Table 5.6, where two bytes, A and B, have

Rnd	Left Half				Right Half				Output XOR				Probability
ΔP	A	B	0	0	0	0	0	0					
1	A	B	0	0	0	0	0	0	0	0	0	0	
2	0	0	0	0	A	B	0	0	0	0	0	0	2^{-14}
ΔC	A	B	0	0	0	0	0	0					

Table 5.6: 2-Round Iterative Characteristic for a Balanced CAST Cipher with $M = 4$, $m = 8$, and $n = 32$

non-zero input XOR values, and result in a zero output XOR value in the XOR table of the round function.

Similarly, for an unbalanced CAST cipher with $M = 2$, $m = 8$, and $n = 32$, suppose that we have discarded all S-boxes with which a 4-round iterative characteristic with a differential probability of 2^{-7} can be constructed. Then, $\lambda_4 = (2^7)^2/2^{32} = 2^{-18}$, and $\mu_4 = (2^8 - 1)^2 \cdot \lambda_4 \approx 2^{-2}$. Therefore, when we randomly generate a round function

combined by two 8×32 S-boxes, the expected number of entries in the XOR table which satisfy $N_E(\Delta X, 0) = 4$ is about 0.25. This entry can be used to construct a 4-round iterative characteristic with a differential probability of 2^{-14} as in Table 5.7. We assume that the best iterative characteristic for an unbalanced CAST cipher with

Rnd	Left Half				Right Half				Output XOR				Probability
ΔP	A	B	0	0	0	0	0	0					
1	A	B	0	0	0	0	0	0	0	0	0	0	
2	0	0	0	0	A	B	0	0	0	0	0	0	
3	0	0	A	B	0	0	0	0	0	0	0	0	
4	0	0	0	0	0	0	A	B	0	0	0	0	2^{-14}
ΔC	A	B	0	0	0	0	0	0					

Table 5.7: 4-Round Iterative Characteristic for an Unbalanced CAST Cipher with $M = 2$, $m = 8$, and $n = 32$

$M = 2$, $m = 8$, and $n = 32$ is a 4-round iterative characteristic with a differential probability of 2^{-14}. Screening a pair of S-boxes to prevent the characteristic from occurrence would be possible although time-consuming since about 2^{30} XOR pairs would have to be examined.

An unbalanced CAST cipher with $M = 1$, $m = 8$, and $n = 32$ has $\lambda_2 = 2^7/2^{32} = 2^{-25}$. and $\mu_2 = (2^8 - 1) \cdot \lambda_2 \approx 2^{-17}$. The expected number is so small that it would be difficult to find an 8-round iterative characteristic as in Table 5.4 with a differential probability of $p = 2^{-7}$ when we randomly generate an 8×32 S-box.

Table 5.8 gives the likelihood of occurrence of iterative characteristics for an unbalanced CAST cipher with $M = 2$, $m = 8$, and $n = 16$. It can be seen that the expected number of entries in the XOR table which can be used to construct 4-round iterative characteristic with a differential probability of 2^{-7} is large, about 1. Even if we discard such S-boxes by using the screening process, the 4-round iterative char-

$NZ(\Delta X)$	v	λ_v	μ_v	p
2	$2^2 \cdot 1$	$2^{-2.36}$	1.27×10^4	2^{-14}
2	$2^2 \cdot 2$	$2^{-5.36}$	1.58×10^3	2^{-13}
2	$2^2 \cdot 3$	$2^{-8.95}$	1.32×10^2	$2^{-12.4}$
2	$2^2 \cdot 4$	$2^{-12.95}$	8.24	2^{-12}
2	$2^2 \cdot 5$	$2^{-17.27}$	4.12×10^{-1}	$2^{-11.7}$
2	$2^2 \cdot 6$	$2^{-21.85}$	1.72×10^{-2}	$2^{-11.42}$
2	$2^2 \cdot 7$	$2^{-26.66}$	6.13×10^{-4}	$2^{-11.19}$
⋮	⋮	⋮	⋮	⋮
1	$2^9 \cdot 1$	2^{-9}	9.94×10^{-1}	2^{-7}
1	$2^9 \cdot 2$	2^{-19}	9.71×10^{-4}	2^{-6}
⋮	⋮	⋮	⋮	⋮

Table 5.8: Likelihood of Occurrence of Iterative Characteristics for an Unbalanced CAST Cipher with $M = 2$, $m = 8$, and $n = 16$

acteristic with a differential probability of 2^{-12} can be easily found, and can not be screened because the likelihood of occurrence is too high, with the expected number of entries of about 8.24. Therefore, we assume that the best iterative characteristic for an unbalanced CAST cipher with $M = 2$, $m = 8$, and $n = 16$ is a 4-round iterative characteristic with a differential probability of 2^{-12}, which is shown in Table 5.9.

Rnd	Left Half				Right Half				Output XOR		Probability
ΔP	A	B	0	0	0	0	0	0			
1	A	B	0	0	0	0	0	0	0	0	
2	0	0	0	0	A	B	0	0	0	0	
3	0	0	A	B	0	0	0	0	0	0	
4	0	0	0	0	0	0	A	B	0	0	2^{-12}
ΔC	A	B	0	0	0	0	0	0			

Table 5.9: 4-Round Iterative Characteristic for an Unbalanced CAST Cipher with $M = 2$, $m = 8$, and $n = 16$

An 8-round iterative characteristic is illustrated in Table 5.10 for an unbalanced CAST cipher with $M = 1$, $m = 8$, and $n = 16$. The cipher has $\lambda_2 = 2^7/2^{16} = 2^{-9}$, and $\mu_2 = (2^8 - 1) \cdot \lambda_2 \approx 0.5$, which is much larger than the one of an unbalanced CAST

Rnd	Left Half				Right Half				Output XOR		Probability
ΔP	A	0	0	0	0	0	0	0			
1	A	0	0	0	0	0	0	0	0	0	
2	0	0	0	0	A	0	0	0	0	0	
3	0	A	0	0	0	0	0	0	0	0	
4	0	0	0	0	0	A	0	0	0	0	
5	0	0	A	0	0	0	0	0	0	0	
6	0	0	0	0	0	0	A	0	0	0	
7	0	0	0	A	0	0	0	0	0	0	
8	0	0	0	0	0	0	0	A	0	0	2^{-7}
ΔC	A	0	0	0	0	0	0	0			

Table 5.10: 8-Round Iterative Characteristic for an Unbalanced CAST Cipher with $M = 1$, $m = 8$, and $n = 16$

cipher with $M = 1$, $m = 8$, and $n = 32$. The screening process is still applicable to discard such S-boxes which can be used to construct an 8-round iterative characteristic since the number of entries in the XOR table of an 8×16 S-box is only $2^8 \cdot 2^7 = 2^{15}$.

Table 5.11 gives the likelihood of occurrence of iterative characteristics for a balanced CAST cipher with $M = 8$, $m = 4$, and $n = 32$. The scenario is quite similar to the one of a balanced CAST cipher with $M = 4$, $m = 8$, and $n = 32$. It is highly unlikely to construct an iterative characteristic with differential probabilities of 2^{-3} or 2^{-6}, for which only one or two of eight S-boxes have an entry corresponding to a non-zero input XOR and a zero output XOR in its XOR table. Although the likelihood of occurrence of an iterative characteristic with a differential probability of 2^{-9} is not low enough, about 0.02, it is not difficult to apply a screening process to discard any such S-boxes. The work load is to check totally $\binom{8}{3} \cdot 2^{12} \cdot (2^3)^3 \approx 2^{26.8}$ output XOR pairs for the round function. However, it is very likely ($\mu_v = 3.38$ for $v = 2^{20}$) to construct an iterative characteristic with a differential probability of 2^{-12}. To discard such S-boxes to prevent this iterative characteristic from occurring by the screening

$NZ(\Delta X)$	v	λ_v	μ_v	p
8	$2^8 \cdot 1$	$2^{-8.01}$	9.97×10^6	2^{-24}
8	$2^8 \cdot 2$	$2^{-17.01}$	1.95×10^4	2^{-23}
8	$2^8 \cdot 3$	$2^{-26.59}$	2.54×10^1	$2^{-22.4}$
8	$2^8 \cdot 4$	$2^{-36.59}$	2.48×10^{-2}	2^{-22}
8	$2^8 \cdot 5$	$2^{-46.91}$	1.93×10^{-5}	$2^{-21.7}$
⋮	⋮	⋮	⋮	⋮
7	$2^{11} \cdot 1$	$2^{-11.00}$	6.67×10^5	2^{-21}
7	$2^{11} \cdot 2$	$2^{-23.00}$	1.63×10^2	2^{-20}
7	$2^{11} \cdot 3$	$2^{-35.59}$	2.65×10^{-2}	$2^{-19.4}$
7	$2^{11} \cdot 4$	$2^{-48.59}$	3.24×10^{-6}	2^{-19}
⋮	⋮	⋮	⋮	⋮
6	$2^{14} \cdot 1$	$2^{-14.00}$	1.95×10^4	2^{-18}
6	$2^{14} \cdot 2$	$2^{-29.00}$	5.94×10^{-1}	2^{-17}
6	$2^{14} \cdot 3$	$2^{-44.59}$	1.21×10^{-5}	$2^{-16.4}$
⋮	⋮	⋮	⋮	⋮
5	$2^{17} \cdot 1$	$2^{-17.00}$	3.24×10^2	2^{-15}
5	$2^{17} \cdot 2$	$2^{-35.00}$	1.24×10^{-3}	2^{-14}
⋮	⋮	⋮	⋮	⋮
4	$2^{20} \cdot 1$	$2^{-20.00}$	3.38	2^{-12}
4	$2^{20} \cdot 2$	$2^{-41.00}$	1.61×10^{-6}	2^{-11}
⋮	⋮	⋮	⋮	⋮
3	$2^{23} \cdot 1$	$2^{-23.00}$	2.25×10^{-2}	2^{-9}
3	$2^{23} \cdot 2$	$2^{-47.00}$	1.34×10^{-9}	2^{-8}
⋮	⋮	⋮	⋮	⋮
2	$2^{26} \cdot 1$	$2^{-26.00}$	9.39×10^{-5}	2^{-6}
⋮	⋮	⋮	⋮	⋮
1	$2^{29} \cdot 1$	$2^{-29.00}$	2.24×10^{-7}	2^{-3}
⋮	⋮	⋮	⋮	⋮

Table 5.11: Likelihood of Occurrence of Iterative Characteristics for an Unbalanced CAST Cipher with $M = 8$, $m = 4$, and $n = 32$

process, the work load is to check totally $\binom{8}{4} \cdot 2^{16} \cdot (2^3)^4 \approx 2^{34.1}$ output XOR pairs for the round function. Therefore, we assume that the best iterative characteristic for a balanced CAST cipher with $M = 8$, $m = 4$, and $n = 32$ is a 2-round iterative characteristic with a differential probability of 2^{-12}, which is a little higher than the one of a balanced CAST cipher with $M = 4$, $m = 8$, and $n = 32$. One of such iterative characteristics is illustrated in Table 5.6, in which the differential probability should be 2^{-12}.

For an unbalanced CAST cipher with $M = 4$, $m = 4$, and $n = 32$, suppose that we have discarded all S-boxes with which a 4-round iterative characteristic with differential probabilities of 2^{-3}, 2^{-6}, or 2^{-9} can be constructed. Then, $\lambda_{16} = 2^{12}/2^{32} = 2^{-20}$, and $\mu_{16} = (2^4 - 1)^4 \cdot \lambda_4 \approx 2^{-4.4}$. Then, when we randomly generate a round function from four 4×32 S-boxes, the expected number of entries in the XOR table which satisfy $N_E(\Delta X, 0) = 16$ is about 0.05. This entry can be used to construct a 4-round iterative characteristic with a differential probability of 2^{-12} as in Table 5.7. We assume that the best iterative characteristic for an unbalanced CAST cipher with $M = 4$, $m = 4$, and $n = 32$ is a 4-round iterative characteristic with a differential probability of 2^{-12}, although it may be possible to screen such S-boxes.

An unbalanced CAST cipher with $M = 2$, $m = 4$, and $n = 32$ has $\lambda_4 = (2^3)^2/2^{32} = 2^{-26}$, and $\mu_4 = (2^4 - 1)^2 \cdot \lambda_2 \approx 2^{-18.2}$. The expected number is so small that it would be difficult to find an 8-round iterative characteristic with a differential probability of 2^{-6} as in Table 5.4 when we randomly generate two 4×32 S-boxes.

We do not consider an unbalanced CAST cipher with $M = 1$, $m = 4$, and $n = 32$

since one 4×32 S-box has at most 2^4 different 32-bit output vectors. Any linear combinations of such 2^4 vectors can not produce all 2^{32} different 32-bit vectors. Hence, the round function is not surjective, and the cipher is trivially breakable using Preneel's basic attack described in Section 2.6.3.

Since the number of output bits of an 8×56 S-box is larger than the one of an 8×32 S-box, it is more difficult to find an 8-round iterative characteristic with a differential probability of 2^{-7} for an unbalanced CAST cipher with $M = 1$, $m = 8$, and $n = 56$.

Finally, a summary of the likelihood of occurrence of iterative characteristics for unbalanced CAST ciphers with some typical parameters are listed in Table 5.12. The difficulty level of which S-boxes could be screened depends on the likelihood of occurrence of iterative characteristics and the work load of screening. We can conclude by noting that iterative characteristics can be prevented for unbalanced CAST ciphers with one 8×56 S-box, with one 8×32 S-box, and with two 4×32 S-boxes.

5.2.2 Pseudo-Iterative Characteristics

By looking at Table 5.12, we can find that it seems difficult for an unbalanced CAST cipher with one 8×32 S-box to have an iterative characteristic constructed by an entry of $N_E(\Delta X, 0)$ in the XOR table of the round function. Even if the XOR table of the round function has an entry of $N_E(\Delta X, 0)$ the likelihood of occurrence of this entry is not so high that we can not discard S-boxes leading to such round functions by the screening process. Therefore, we believe that differential cryptanalysis of the unbalanced CAST cipher whose round function is constructed by one 8×32 S-box by

S-box		Differential	Notes
$m \times n$	M	Probability	
8×56	1	2^{-7} (8 Rounds)	Likelihood: $\mu = 4.53 \times 10^{-13}$ S-boxes could be easily screened
8×32	1	2^{-7} (8 Rounds)	Likelihood: $\mu = 7.63 \times 10^{-6}$ S-boxes could be easily screened
	2	2^{-14} (4 Rounds)	Likelihood: $\mu = 0.25$ S-boxes could be screened
	4	2^{-14} (2 Rounds)	Likelihood: $\mu = 1.49$ S-boxes might be screened with difficulty
8×16	1	2^{-7} (8 Rounds)	Likelihood: $\mu = 0.50$ S-boxes could be screened
	2	2^{-12} (4 Rounds)	Likelihood: $\mu = 8.28$ S-boxes might be screened with difficulty
4×32	2	2^{-6} (8 Rounds)	Likelihood: $\mu = 3.32 \times 10^{-6}$ S-boxes could be easily screened
	4	2^{-12} (4 Rounds)	Likelihood: $\mu = 0.05$ S-boxes could be screened
	8	2^{-12} (2 Rounds)	Likelihood: $\mu = 3.38$ S-boxes might be screened with difficulty

Table 5.12: Summary of Likelihood of Occurrence of Iterative Characteristics

using the iterative characteristic described in Section 5.2.1 is not a real threat. In this section, we investigate the probability that there exist other kinds of characteristics similar in nature to iterative characteristics, but not strictly iterative. We refer to these characteristics as pseudo-iterative characteristics.

Table 5.13 displays an 8-round pseudo-iterative characteristic for an unbalanced CAST cipher with $M = 1$, $m = 8$, and $n = 32$, where any letter represents a non-zero byte of XOR values, $A_1 = A \oplus a_1$, and $B_1 = B \oplus b_1$. The differential probability of this characteristic is 2^{-14}. In terms of the characteristic, the XOR table of the round function must have entries of $N_E(B, 000a_1)$ and $N_E(A_1, b_1000)$ at the same time. Furthermore, if there are entries of $N_E(B_1, 000a_1)$ and $N_E(A, b_1000)$ in the

Rnd	Left Half				Right Half				Output XOR				Probability
ΔP	A	0	0	0	B	0	0	0					
1	A	0	0	0	B	0	0	0	0	0	0	0	
2	0	B	0	0	A	0	0	0	0	0	0	0	
3	0	A	0	0	0	B	0	0	0	0	0	0	
4	0	0	B	0	0	A	0	0	0	0	0	0	
5	0	0	A	0	0	0	B	0	0	0	0	0	
6	0	0	0	B	0	0	A	0	0	0	0	0	
7	0	0	0	A	0	0	0	B	0	0	0	a_1	2^{-7}
8	B	0	0	0	0	0	0	A_1	b_1	0	0	0	2^{-7}
ΔC	A_1	0	0	0	B_1	0	0	0					

Table 5.13: First 8-Round Pseduo-Iterative Characteristic for an Unbalanced CAST Cipher with $M = 1$, $m = 8$, and $n = 32$

XOR table, the characteristic can be extended to a 16-round iterative characteristic with a differential probability of 2^{-28}.

The XOR table has totally $(2^8 - 1) \cdot 2^7 \approx 2^{15}$ non-zero entries. Suppose that they are randomly generated and independent. Then the probability that there is at least one entry of the form of $X000$ in the XOR table is $1 - (1 - \frac{2^8}{2^{32}})^{2^{15}} \approx 2^{-9}$, where X is a non-zero byte. Similarly, the probability that there is at least one entry of the form of $000X$ in the XOR table is about 2^{-9}. The probability that there is at least one pair of entries of the forms of $X000$ and $000X$ in the XOR table is approximately $(2^{-9})^2 = 2^{-18}$. Furthermore, the likelihood of the actual entries required for the pseudo-iterative characteristics is even smaller. Hence, it is extremely unlikely for the cipher to have the characteristic of Table 5.13 or the resulting 16-round iterative characteristic.

Table 5.14 displays another 8-round pseudo-iterative characteristic for an unbalanced CAST cipher with $M = 1$, $m = 8$, and $n = 32$. The differential probability of this characteristic is 2^{-21}. The characteristic requires that the XOR table of the

Rnd	Left Half				Right Half				Output XOR				Probability
ΔP	A	B	0	0	C	0	0	0					
1	A	B	0	0	C	0	0	0	0	0	0	0	
2	0	C	0	0	A	B	0	0	0	0	0	0	
3	0	A	B	0	0	C	0	0	0	0	0	0	
4	0	0	C	0	0	A	B	0	0	0	0	0	
5	0	0	A	B	0	0	C	0	0	0	0	0	
6	0	0	0	C	0	0	A	B	0	0	0	c_1	2^{-7}
7	B	0	0	A	0	0	0	C_1	b_1	0	0	a_1	2^{-7}
8	C_1	0	0	0	B_1	0	0	A_1	c_2	0	0	0	2^{-7}
ΔC	A_1	B_1	0	0	C_2	0	0	0					

Table 5.14: Second 8-Round Pseudo-Iterative Characteristic for an Unbalanced CAST Cipher with $M = 1$, $m = 8$, and $n = 32$

round function must have entries of the forms of $000X$, $X00X$, and $X000$ at the same time. The probability that the XOR table has these entries is approximately $2^{-9} \cdot 2^{-1} \cdot 2^{-9} = 2^{-19}$, which is still very small. Further, the probability of actual entries required for an pseudo-iterative characteristic is much smaller.

Table 5.15 displays third 8-round pseudo-iterative characteristic for an unbalanced CAST cipher with $M = 1$, $m = 8$, and $n = 32$. The differential probability of this

Rnd	Left Half				Right Half				Output XOR				Probability
ΔP	A	B	0	0	C	D	0	0					
1	A	B	0	0	C	D	0	0	0	0	0	0	
2	0	C	D	0	A	B	0	0	0	0	0	0	
3	0	A	B	0	0	C	D	0	0	0	0	0	
4	0	0	C	D	0	A	B	0	0	0	0	0	
5	0	0	A	B	0	0	C	D	0	0	a_1	b_1	2^{-7}
6	D	0	0	C	0	0	A_1	B_1	d_1	0	0	c_1	2^{-7}
7	B_1	0	0	A_1	D_1	0	0	C_1	b_2	0	0	a_2	2^{-7}
8	C_1	D_1	0	0	B_2	0	0	A_2	c_2	d_2	0	0	2^{-7}
ΔC	A_2	B_2	0	0	C_2	D_2	0	0					

Table 5.15: Third 8-Round Pseudo-Iterative Characteristic for an Unbalanced CAST Cipher with $M = 1$, $m = 8$, and $n = 32$

characteristic is 2^{-28} for eight rounds. The characteristic requires that the XOR table

of the round function must have two entries of the forms of $00XX$ and $XX00$ and two entries of the form of $X00X$. The probability that the XOR table has these entries is approximately $(2^{-1})^4 = 2^{-4}$. The probability of occurrence of the pseudo-iterative characteristic in Table 5.15 is much higher than the previous two tables, although it is still an upper bound and the real probability will be lower since it requires specific relationships between the non-zero bytes of the output XOR of the round function. However, since the likelihood of occurrence of Table 5.15 is not low enough, we conclude that an 8-round pseudo-iterative characteristic with a differential probability of 2^{-28} for an unbalanced CAST cipher with one 8×32 S-box is possible.

The same analysis can be applied to an unbalanced CAST cipher with two 4×32 S-boxes. As a result, we assume that the cipher has an 8-round pseudo-iterative characteristic with a differential probability of $(2^{-6})^4 = 2^{-24}$.

5.2.3 Effectiveness of Iterative Characteristics

In general, by applying the best iterative or the pseudo-iterative characteristic and using an $(R - N/l)$-round attack on an R-round unbalanced CAST cipher, the probability of a right pair occurring will be

$$P_{\Omega_{(R-N/l)}} = p^{\frac{(R-N/l)}{N/l}} \tag{5.8}$$

where p is a differential probability of a best N/l-round iterative or a pseudo-iterative characteristic. The differential attack needs η right pairs. The number of chosen plaintexts required to uniquely identify a right subkey is given by $N_D = \eta / P_{\Omega_{(R-N/l)}}$. Based on the iterative characteristics given in Table 5.12 and the ones assumed in Section 5.2.2, a summary of the differential cryptanalysis is shown in Table 5.16.

S-box		Number of Rounds	Probability of Characteristic	Plaintexts Needed
$m \times n$	M			
8×56	1	–	–	–
8×32	1	24	2^{-56}	$\eta \cdot 2^{56}$
		32	2^{-84}	$\eta \cdot 2^{84}$
	2	20	2^{-56}	$\eta \cdot 2^{56}$
		24	2^{-70}	$\eta \cdot 2^{70}$
	4	10	2^{-56}	$\eta \cdot 2^{56}$
		12	2^{-70}	$\eta \cdot 2^{70}$
8×16	1	80	2^{-63}	$\eta \cdot 2^{63}$
		88	2^{-70}	$\eta \cdot 2^{70}$
	2	24	2^{-60}	$\eta \cdot 2^{60}$
		28	2^{-72}	$\eta \cdot 2^{72}$
4×32	2	24	2^{-48}	$\eta \cdot 2^{48}$
		32	2^{-72}	$\eta \cdot 2^{72}$
	4	24	2^{-60}	$\eta \cdot 2^{60}$
		28	2^{-72}	$\eta \cdot 2^{72}$
	8	12	2^{-60}	$\eta \cdot 2^{60}$
		14	2^{-72}	$\eta \cdot 2^{72}$

Table 5.16: Summary of Differential Cryptanalysis Based on the Best Iterative and the Pseudo-Iterative Characteristics

We do not include the case of an unbalanced CAST cipher with one 8×56 S-box since the likelihood of occurrence of an 8-round iterative characteristic with a differential probability of 2^{-7} is so small that it is extremely unlikely. On the other hand, we can not find an 8-round pseudo-iterative characteristic with a reasonable differential probability in our analysis.

5.3 Biham's Characteristics

There is another approach for differential cryptanalysis of unbalanced CAST ciphers, which is somewhat different from the generic differential cryptanalysis considered in Section 5.1 and Section 5.2. This approach was introduced by Biham [5]. Biham

applies the attack to Khafre[1] [21], which can be seen as an unbalanced CAST cipher with $M = 1$, $m = 8$, and $n = 32$. In this section, we extend Biham's attack to our general class of unbalanced CAST ciphers. It appears to be another practical way to differentially attack the unbalanced CAST ciphers with reduced rounds.

The main idea of differential cryptanalysis of the unbalanced CAST ciphers is based on the fact that the number of input bits of an S-box is less than the number of output bits. From the analysis results of Section 5.1, it can be seen that for an S-box with $n \gg m$, the maximum entry value of the output XORs is practically 2 for a particular input XOR. Furthermore, since there are totally $2^m \cdot 2^{m-1}$ input XOR pairs in the XOR table of an S-box, which is usually less than 2^n possible output XORs, only about $\frac{2^{2m-1}}{2^n}$ output XORs will occur. For example, an 8×32 S-box has about 2^{15} input XORs which is 2^{-17} of 2^{32} possible output XORs. An 8×16 S-box has 2^{15} input XORs and, hence, less than about half of 2^{16} possible output XORs are used.

Table 5.17 illustrates one of Biham's characteristics used to attack a 16-round unbalanced CAST cipher with $M = 1$, $m = 8$, and $n = 32$. We categorize it as Biham's Type I characteristic. The rotation and swapping operation are defined in Section 4.3.3. Each value 0 describes a zero byte XOR value. Each letter denotes a non-zero byte XOR value. The exact non-zero XOR values may vary for different right pairs. The superscript † means that the byte of the output XOR must be equal to the corresponding byte of the left half such that the input XOR of the S-box in the next round will be zero. Each occurrence of † causes a reduction of the probability

[1] Khafre uses a different S-box after every eight rounds.

Rnd	Left Half				Right Half				Output XOR				Probability
ΔP	0	A	0	0	0	0	0	0					
1	0	A	0	0	0	0	0	0	0	0	0	0	
2	0	0	0	0	0	A	0	0	0	0	0	0	
3	0	0	A	0	0	0	0	0	0	0	0	0	
4	0	0	0	0	0	0	A	0	0	0	0	0	
5	0	0	0	A	0	0	0	0	0	0	0	0	
6	0	0	0	0	0	0	0	A	w_1	x_1	y_1	z_1	
7	A	0	0	0	w_1	x_1	y_1	z_1	B	C	D	E	
8	z_1	w_1	x_1	y_1	F	C	D	E	z_2	w_2	x_2	$y_1^\dagger$	$p = 2^{-8}$
9	E	F	C	D	z_3	w_3	x_3	0	0	0	0	0	
10	0	z_3	w_3	x_3	E	F	C	D	y_4	z_4	w_4	$x_3^\dagger$	$p = 2^{-8}$
11	D	E	F	C	y_5	z_5	w_5	0	0	0	0	0	
12	0	y_5	z_5	w_5	D	E	F	C	x_6	y_6	z_6	$w_5^\dagger$	$p = 2^{-8}$
13	C	D	E	F	x_7	y_7	z_7	0	0	0	0	0	
14	0	x_7	y_7	z_7	C	D	E	F	w_8	x_8	y_8	$z_7^\dagger$	$p = 2^{-8}$
15	F	C	D	E	w_9	x_9	y_9	0	0	0	0	0	
16	0	w_9	x_9	y_9	F	C	D	E	z_{10}	w_{10}	x_{10}	y_{10}	
ΔC	E	F	C	D	z_{11}	w_{11}	x_{11}	y_{11}					

Table 5.17: Biham's Type I Characteristic of a 16-Round Unbalanced CAST Cipher with $M = 1$, $m = 8$, and $n = 32$

of the characteristic by $1/255 \approx 2^{-8}$. Therefore, the probability of the characteristic is about 2^{-32}.

From the characteristic, we can see that the output XOR of the 7-th round can be easily extracted by XORing the left half of the plaintext XOR with the left half of the ciphertext XOR and rotating the result by eight bits. This happens because the 7-th round is the only odd round whose output XOR is not zero. Since there are only about 2^{15} possible input XOR pairs for the S-box, there are at most 2^{15} possible output XOR values for $BCDE$, which is only 2^{-17} of 2^{32} all possible XOR values. As a result, most of wrong pairs can be easily discarded by comparing the $BCDE$ to all possible output XOR values of the S-box. If a right pair occurs, we know the

actual values of the input XOR E to the S-box in the 16-th round and one byte of the output XOR, $z_{10} = z_{11}$. Since there are only 2^7 output XORs for that input XOR E, by looking up the XOR table, the output XOR of the last round function $z_{10}w_{10}x_{10}y_{10}$ can be decided with a very high probability. Using the method described in Section 3.1.1, we can count the subkey bits used in the 16-th round.

This attack needs η right pairs obtained from a pool of about $\eta \cdot 2^{32}$ pairs which are formed by $\eta \cdot 2^{33}$ chosen plaintexts. This number of plaintexts can be drastically reduced by using a compact structure of 2^8 encryptions which generates 2^{15} pairs [5]. Since the plaintext XOR ΔP has one degree of freedom, A, which can have $2^8 - 1$ possible values, the compact structure chooses all 2^8 possible values for that byte and a constant random value for remaining seven bytes of the plaintexts, and encrypts all plaintexts. Therefore, the attack needs $\eta \cdot 2^{32} \cdot 2^8/2^{15} = \eta \cdot 2^{25}$ encryptions.

An even better characteristic with a probability of 2^{-16} is shown in Table 5.18. We categorize it as Biham's Type II characteristic. The left half of the ciphertext XOR is the XOR sum of the output XORs of the 9-th (by one byte rotation) and 11-th rounds, as well as the non-zero plaintext XOR byte A. This XOR sum has at most 2^{30} possible values, and this can be used to discard the wrong pairs. Using the same compact structure, $\eta \cdot 2^{16} \cdot 2^8/2^{15} = \eta \cdot 2^9$ encryptions (or plaintexts) are needed. It should be noticed that although Type II characteristic has a larger probability p, it also has a larger γ since wrong pairs can not be discarded efficiently and more counting work has to be done. So the *SNR*s of two counting schemes could be similar, and we believe that in both cases η is small. Although this implies that the Type II characteristic

Rnd	Left Half				Right Half				Output XOR				Probability
ΔP	A	0	0	0	0	0	0	0					
⋮		⋮				⋮		⋮		⋮			
8	0	0	0	0	0	0	0	A	w_1	x_1	y_1	z_1	
9	A	0	0	0	w_1	x_1	y_1	z_1	B	C	D	E	
10	z_1	w_1	x_1	y_1	F	C	D	E	z_2	w_2	x_2	y_2	
11	E	F	C	D	z_3	w_3	x_3	y_3	H	I	J	K	
12	y_3	z_3	w_3	x_3	L	M	N	O	y_4	z_4	w_4	$x_3^\dagger$	$p = 2^{-8}$
13	O	L	M	N	y_5	z_5	w_5	0	0	0	0	0	
14	0	y_5	z_5	w_5	O	L	M	N	x_6	y_6	z_6	$w_5^\dagger$	$p = 2^{-8}$
15	N	O	L	M	x_7	y_7	z_7	0	0	0	0	0	
16	0	x_7	y_7	z_7	N	O	L	M	w_8	x_8	y_8	z_8	
ΔC	M	N	O	L	w_9	x_9	y_9	z_9					

Table 5.18: Biham's Type II Characteristic of a 16-Round Unbalanced CAST Cipher with $M = 1$, $m = 8$, and $n = 32$

is more useful in an attack, it should be noted that for every plaintext difference, a lot more work is necessary in the counting scheme for the Type II characteristic than for the Type I characteristic. Both characteristics can be extended to characteristics with more than sixteen rounds. However, every additional eight rounds will have a 2^{-32} fixed reduction of the probability. Thus the characteristics become quickly impractical to be applied for many more rounds.

Table 5.19 shows Biham's Type I characteristic used to attack an 8-round unbalanced CAST cipher with $M = 2$, $m = 8$, and $n = 32$. Table 5.20 shows Biham's Type II characteristic. Both characteristics have a 2^{-32} fixed reduction of the differential probability for every additional four rounds. Although Type I characteristic has lower differential probability than Type II, $DEFG$ in Type I has only about 2^{15} possible XOR values, but $CDEF$ in Type II has about 2^{30}. Therefore, Type I can be used to discard wrong pairs more efficiently than Type II. Using a compact

Rnd	Left Half				Right Half				Output XOR				Probability
ΔP	A	B	0	0	C	0	0	0					
1	A	B	0	0	C	0	0	0	0	0	0	0	
2	0	0	C	0	A	B	0	0	0	0	0	0	
3	0	0	A	B	0	0	C	0	D	E	F	G	
4	C	0	0	0	D	E	H	I	w_1	x_1	$0^\dagger$	$0^\dagger$	$p = 2^{-16}$
5	H	I	D	E	w_2	x_2	0	0	0	0	0	0	
6	0	0	w_2	x_2	H	I	D	E	y_1	z_1	$w_2^\dagger$	$x_2^\dagger$	$p = 2^{-16}$
7	D	E	H	I	y_2	z_2	0	0	0	0	0	0	
8	0	0	y_2	z_2	D	E	H	I	w_3	x_3	y_3	z_3	
ΔC	H	I	D	E	w_4	x_4	y_4	z_4					

Table 5.19: Biham's Type I Characteristic of an 8-Round Unbalanced CAST Cipher with $M = 2$, $m = 8$, and $n = 32$

Rnd	Left Half				Right Half				Output XOR				Probability
ΔP	A	B	0	0	0	0	0	0					
⋮			⋮			⋮		⋮			⋮		
4	0	0	0	0	0	0	A	B	w_1	x_1	y_1	z_1	
5	A	B	0	0	w_1	x_1	y_1	z_1	C	D	E	F	
6	y_1	z_1	w_1	x_1	G	H	E	F	y_2	z_2	$w_1^\dagger$	$x_1^\dagger$	$p = 2^{-16}$
7	E	F	G	H	y_3	z_3	0	0	0	0	0	0	
8	0	0	y_3	z_3	E	F	G	H	w_4	x_4	y_4	z_4	
ΔC	G	H	E	F	w_5	x_5	y_5	z_5					

Table 5.20: Biham's Type II Characteristic of an 8-Round Unbalanced CAST Cipher with $M = 2$, $m = 8$, and $n = 32$

structure, $\eta \cdot 2^9$ encryptions are needed to attack the 8-round cipher with Type II characteristic, which is similar to the amount of encryptions required to attack a 16-round unbalanced CAST cipher with $M = 1$, $m = 8$, and $n = 32$ using Type II characteristic.

Since an unbalanced CAST cipher with $M = 2$, $m = 8$, and $n = 32$ has the XOR table of the round function twice as large as an unbalanced CAST cipher with $M = 1$, $m = 8$, and $n = 32$, the differential cryptanalysis of the cipher with $M = 2$ requires more computing time to identify the right and wrong pairs. Therefore, the cipher

with $M = 2$ is a little stronger than the one with $M = 1$ if Equation 4.2 holds for both ciphers.

Table 5.21 is Biham's Type II characteristic used to attack a 16-round unbalanced CAST cipher with $M = 1$, $m = 8$, and $n = 16$, and has a probability of 2^{-16}. The

Rnd	Left Half				Right Half				Output XOR		Probability
ΔP	A	0	0	0	0	0	0	0			
⋮		⋮				⋮		⋮	⋮		
8	0	0	0	0	0	0	0	A	y_1	z_1	
9	A	0	0	0	0	0	y_1	z_1	B	C	
10	z_1	0	0	y_1	A	0	B	C	x_1	$y_1^\dagger$	$p = 2^{-8}$
11	C	A	0	B	z_1	0	x_1	0	0	0	
12	0	z_1	0	x_1	C	A	0	B	y_2	x_2	
13	B	C	A	0	0	z_1	y_2	x_3	D	E	
14	x_3	0	z_1	y_2	B	C	F	E	z_2	$y_2^\dagger$	$p = 2^{-8}$
15	E	B	C	F	x_3	0	z_3	0	0	0	
16	0	x_3	0	z_3	E	B	C	F	y_3	z_4	
ΔC	F	E	B	C	0	x_3	y_3	z_5			

Table 5.21: Biham's Type II Characteristic of a 16-Round Unbalanced CAST Cipher with $M = 1$, $m = 8$, and $n = 16$

characteristic also has a 2^{-32} fixed reduction of the differential probability for every additional eight rounds. By looking at the Table 5.21, we find that the output XORs of the 9-th and the 13-th rounds, BC and DE, have about 2^{15} possible XOR values, and $FEBC$ has at most 2^{30} possible XOR values. Furthermore, we notice that there is a zero-byte in the ciphertext XOR. It is helpful to discard wrong pairs and increase its *SNR*. As a result, we believe that an unbalanced CAST cipher with $M = 1$, $m = 8$, and $n = 16$ is more susceptible to Biham's attack than an unbalanced CAST cipher with $M = 1$, $m = 8$, and $n = 32$, based on the same number of rounds. Using a compact structure, it requires $\eta \cdot 2^9$ encryptions to attack the 16-round cipher with

the characteristic of Table 5.21.

An unbalanced CAST cipher with $M = 2$, $m = 4$, and $n = 32$ has the same characteristics of Table 5.17 and Table 5.18. Also, an unbalanced CAST cipher with $M = 4$, $m = 4$, and $n = 32$ has the same characteristics of Table 5.19 and Table 5.20. However, since the round functions of the ciphers are constructed by small 4×32 S-boxes, and their possible output XOR values are less than the corresponding round functions combined by large S-boxes, the characteristics have high *SNRs*. For example, although an unbalanced CAST cipher with four 4×32 S-boxes has the same characteristic probability as an unbalanced CAST cipher with two 8×32 S-boxes, the possible output XOR values of the round function are about $2^{16} \cdot (2^3)^4 = 2^{28}$ for four 4×32 S-boxes, and about $2^{16} \cdot (2^7)^2 = 2^{30}$ for two 8×32 S-boxes. It is easier for the characteristic constructed by small S-boxes to discard wrong pairs. Therefore, the round function constructed by small S-boxes (e.g. 4×32) is weaker than the corresponding one combined by large S-boxes (e.g. 8×32).

Biham's attack is more difficult to apply to unbalanced CAST ciphers whose output XORs of the round functions can take all possible values of 2^n, because we can not identify the right and wrong pairs directly. However, because output XORs of the round function are not uniformly distributed, a statistical attack may be still valid, although the implementation of the cryptanalysis is significantly more complicated. Such unbalanced CAST ciphers have the round functions constructed by eight 4×32 S-boxes, two 8×16 S-boxes, three and four 8×32 S-boxes.

Biham's attack is not applicable for unbalanced CAST ciphers with $M = 1$, $m = 8$,

and $n = 56$, and $M = 2$, $m = 8$, and $n = 48$, since both ciphers have no swapping and the output bytes of the round functions are XORed with all data bytes except the bytes fed into the round functions.

In general, Biham's Type I characteristic of a $2N/l$-round unbalanced CAST cipher has a differential probability of 2^{-32}, and Type II characteristic of the $2N/l$-round unbalanced CAST cipher has a differential probability of 2^{-16}. On the other hand, Type I characteristics can be used to discard wrong pairs more easily than Type II characteristics. Therefore, both types of characteristics could have similar *SNR*s although the work involved in the counting scheme of Type II characteristics is much more extensive. Both characteristics have a 2^{-32} fixed reduction of the differential probability for every additional N/l rounds.

A summary of the differential cryptanalysis results of unbalanced CAST ciphers based on Biham's Type II characteristics is given in Table 5.22.

5.4 Conclusion

In this chapter, we have extended the method of analyzing the distribution of entries in the XOR table to unbalanced CAST ciphers, and statistically analyzed the likelihood of occurrence of N/l-round iterative characteristics with reasonable differential probabilities for unbalanced CAST ciphers. As a result, it is possible for an unbalanced CAST cipher with two 8×32 S-boxes to have a 4-round iterative characteristic with a differential probability of 2^{-14}. Since it is difficult for the unbalanced CAST cipher with one 8×32 S-box to have an 8-round iterative characteristic with a differential probability of 2^{-7}, we have examined pseudo-iterative characteristics and have

S-box		Number of	Probability of	Plaintexts	*SNR*
$m \times n$	M	Rounds	Characteristic	Needed	
8×56	1	–	–	–	–
8×32	1	16	2^{-16}	$\eta \cdot 2^{9}$	medium
		24	2^{-48}	$\eta \cdot 2^{41}$	medium
		32	2^{-80}	$\eta \cdot 2^{73}$	medium
	2	8	2^{-16}	$\eta \cdot 2^{9}$	low
		12	2^{-48}	$\eta \cdot 2^{41}$	low
		16	2^{-80}	$\eta \cdot 2^{73}$	low
	4	–	–	–	–
8×16	1	16	2^{-16}	$\eta \cdot 2^{9}$	high
		24	2^{-48}	$\eta \cdot 2^{41}$	high
		32	2^{-80}	$\eta \cdot 2^{73}$	high
	2	–	–	–	–
4×32	2	16	2^{-16}	$\eta \cdot 2^{9}$	highest
		24	2^{-48}	$\eta \cdot 2^{41}$	highest
		32	2^{-80}	$\eta \cdot 2^{73}$	highest
	4	8	2^{-16}	$\eta \cdot 2^{9}$	very high
		12	2^{-48}	$\eta \cdot 2^{41}$	very high
		16	2^{-80}	$\eta \cdot 2^{73}$	very high
	8	–	–	–	–

Table 5.22: Result of Differential Cryptanalysis Based on Biham's Type II Characteristics

found that an 8-round iterative characteristic with a differential probability of 2^{-28} is reasonable one.

As well, we have done differential cryptanalysis of unbalanced CAST ciphers with two types of Biham's characteristics. In general, Biham's Type II characteristic of a $2N/l$-round unbalanced CAST cipher has a differential probability of 2^{-16}, and a 2^{-32} fixed reduction of differential probability for every additional N/l rounds. Our analysis shows that a 32-round unbalanced CAST cipher with one 8×32 S-box and a 16-round unbalanced CAST cipher with two 8×32 S-boxes are resistant to differential cryptanalysis, based on Biham's Type II characteristics.

In particular, by combining the conclusions of the iterative characteristics and Biham's characteristics, the 32-round unbalanced CAST cipher with one 8×32 S-box which is equivalent to the 8-round original CAST cipher according to Equation 4.2 and the 24-round unbalanced CAST cipher with two 8×32 S-boxes which is equivalent to the 12-round original CAST cipher require more chosen plaintexts than the 2^{64} possible plaintexts and hence appear to be resistant to both forms of differential cryptanalysis.

Chapter 6

Linear Cryptanalysis of Unbalanced CAST Ciphers

In this chapter, we examine the resistance of unbalanced CAST ciphers to linear cryptanalysis. As in the previous chapter, we first analyze the nonlinearity of S-boxes and the linear cryptanalysis of the whole cipher structure. Again, we enumerate a set of typical parameters.

6.1 Nonlinearity of S-boxes

Heys and Tavares [14] have discussed the likelihood of randomly selecting highly nonlinear S-boxes of CAST. Given an affine m-bit boolean function, g, the probability of randomly selecting another m-bit boolean function, f, so that the hamming distance between f and g, $d(f,g) = 2^{m-2}$, is given by

$$P(d(f,g) = 2^{m-2}) = \left(\begin{array}{c} 2^m \\ 2^{m-2} \end{array} \right) / 2^{2^m}. \tag{6.1}$$

For reasonable values of m, the probability of $d(f,g) < 2^{m-2}$ is bounded as $P(d(f,g) < 2^{m-2}) < P(d(f,g) = 2^{m-2})$. Since any two affine m-bit boolean functions are separated by a distance of 2^{m-1} or 2^m, and there are 2^{m+1} m-bit affine boolean functions,

then

$$P(N(f) < 2^{m-2}) < \rho = 2^{m+1} \cdot P(d(f,g) = 2^{m-2}). \tag{6.2}$$

Assuming that 2^n m-bit boolean functions generated as a linear combination of the n m-bit boolean functions are selected from a set of randomly generated m-bit boolean functions, then the expected number of m-bit boolean functions which have a nonlinearity less than 2^{m-2} for an $m \times n$ S-box is given by

$$E(\#\{f|N(f) < 2^{m-2}\}) < 2^n \cdot \rho. \tag{6.3}$$

In cases where the expected value is very small, the probability that the nonlinearity of the S-box is less than 2^{m-2} is also given approximately by

$$P(N(S) < 2^{m-2}) < 2^n \cdot \rho. \tag{6.4}$$

In the following sections, we discuss two issues related to nonlinearity of an S-box.

6.1.1 Effect of Output Size of an S-box on Its Nonlinearity

In Chapter 5, we have seen that the probability of all the entries in the XOR table of an S-box can be reduced by increasing the number of output bits of the S-box. It is very likely that the entries in the XOR table will have only values 0 and 2, if the number of output bits of the S-box is large enough. However, linear cryptanalysis puts a restriction on n, the size of the output of the S-box. From Equation 6.1 and Equation 6.2, the probability that a randomly generated m-bit boolean function is one of 2^{m+1} affine boolean functions is equal to $P_A = 2^{m+1}/2^{2^m} = 2^{m+1-2^m}$. The expected number of m-bit boolean functions which are affine boolean functions for an

S-box	$N_{\min}$	$E(\#\{f \mid N(f) < N_{\min}\})$
8×16	64	$2^{-27.4}$
	⋮	⋮
8×32	64	$2^{-11.4}$
	⋮	⋮
8×56	64	$2^{12.6}$
	62	$2^{9.4}$
	60	$2^{6.0}$
	58	$2^{2.6}$
	56	$2^{-1.0}$
	54	$2^{-4.8}$
	52	$2^{-8.6}$
	50	$2^{-12.6}$
	⋮	⋮

Table 6.1: Expected Number of Boolean Functions with $N(f) < N_{\min}$

$m \times n$ S-box is equal to $2^n \cdot P_A = 2^{(n-2^m+m+1)}$. If $n \geq 2^m - m - 1$, an $m \times n$ S-box is expected to have an m-bit boolean function linearly derived from the n m-bit boolean functions which is equal to one of 2^{m+1} affine boolean functions. This might allow the cipher to be trivially broken. Therefore, it is potentially dangerous to make the cipher secure by increasing the number of output bits of the S-box with no limitation. For example, since a 4×32 S-box meets the relationship of $n > 2^m - m - 1$, the S-box should have some affine boolean functions. This is why we can now rule out unbalanced CAST ciphers constructed by 4×32 S-boxes from our analysis of the linear attack. Such S-boxes should be avoided because of the potential risk of linear cryptanalysis.

Table 6.1 lists the expected number of m-bit boolean functions that have a non-linearity less than $N_{\min}$ for different sizes of S-boxes, based on an analysis similar to Equation 6.3. It is obvious that the probability that a boolean function has a

lower nonlinearity becomes high with the increasing of n for $8 \times n$ S-boxes. It is very unlikely that an 8×32 S-box has an 8-bit boolean function whose nonlinearity is less than 64. On the other hand, it is very likely for an 8×56 S-box to have an 8-bit boolean function whose nonlinearity is less than 64, but unlikely for the S-box to have an 8-bit boolean function whose nonlinearity is less than 50.

6.1.2 Discussion of Assumptions Used in the Analysis

The above analysis assumes that the 2^n m-bit boolean functions linearly derived from n m-bit boolean functions are randomly selected. We conjecture that this assumption is quite reasonable, and in this section we provide an experimental justification for the assumption.

In our analysis, we have assumed that all n m-bit boolean functions of an S-box are randomly generated, and each single bit in an m-bit boolean function's truth table has an equal chance to be 0 and 1. Due to the independence between n m-bit boolean functions, each bit of a linear XOR sum of the n m-bit boolean functions will still have an equal chance to be 0 and 1. As a result, we can conclude that all 2^n m-bit boolean functions are randomly selected, although possibly correlated. Consider that the XOR sum of two m-bit linear boolean functions is still linear; however, the probability that a randomly generated m-bit boolean function is linear is very small, about $2^{m+1}/2^{2^m}$. For example, the probability that a randomly generated 8-bit boolean function is linear is $2^9/2^{256} = 2^{-247}$, which is too small to be of concern. Hence, even though the 2^n m-bit boolean functions of an S-box can be correlated, the correlation is probably weak.

In general, the value of n is often too large to exactly compute the nonlinearity of an $m \times n$ S-box. For example, an 8×32 S-box must have 2^{32} or about four billion linear combinations of 32 8-bit boolean functions examined. Since we do not know the theoretical nonlinearity distribution of an $m \times n$ S-box, we experimentally examine the distribution by sampling a large subset of linear combinations. In other words, we examine the nonlinearity distribution by randomly picking a subset of n m-bit boolean functions, XORing them and deriving its minimum distance to the set of m-bit affine boolean functions. The nonlinearity distribution of an $8\times$ 32 S-box with 2^{22} such selections is illustrated in Figure 6.1.

For comparison, we experimentally examine the distribution of the nearest distance to the set of m-bit affine boolean functions for randomly generated m-bit boolean functions. Figure 6.2 gives the nonlinearity distribution of randomly generated 8-bit boolean functions by averaging four sets of 2^{22} randomly generated 8-bit boolean functions. Note that two distributions of Figure 6.1 and Figure 6.2 are similar in shape.

In order to test how close these two distributions are, we use the chi-square test [9], a goodness-of-fit test, on our experimental data. The chi-square variable can be used to test whether a set of observed frequencies and a set of expected frequencies are close enough so that we can conclude that they come from the same probability distribution. The expected frequencies can be thought of as the average number of values expected to fall in each category, based on some theoretical probability distribution. The observed frequencies can be thought of as a sample of values from

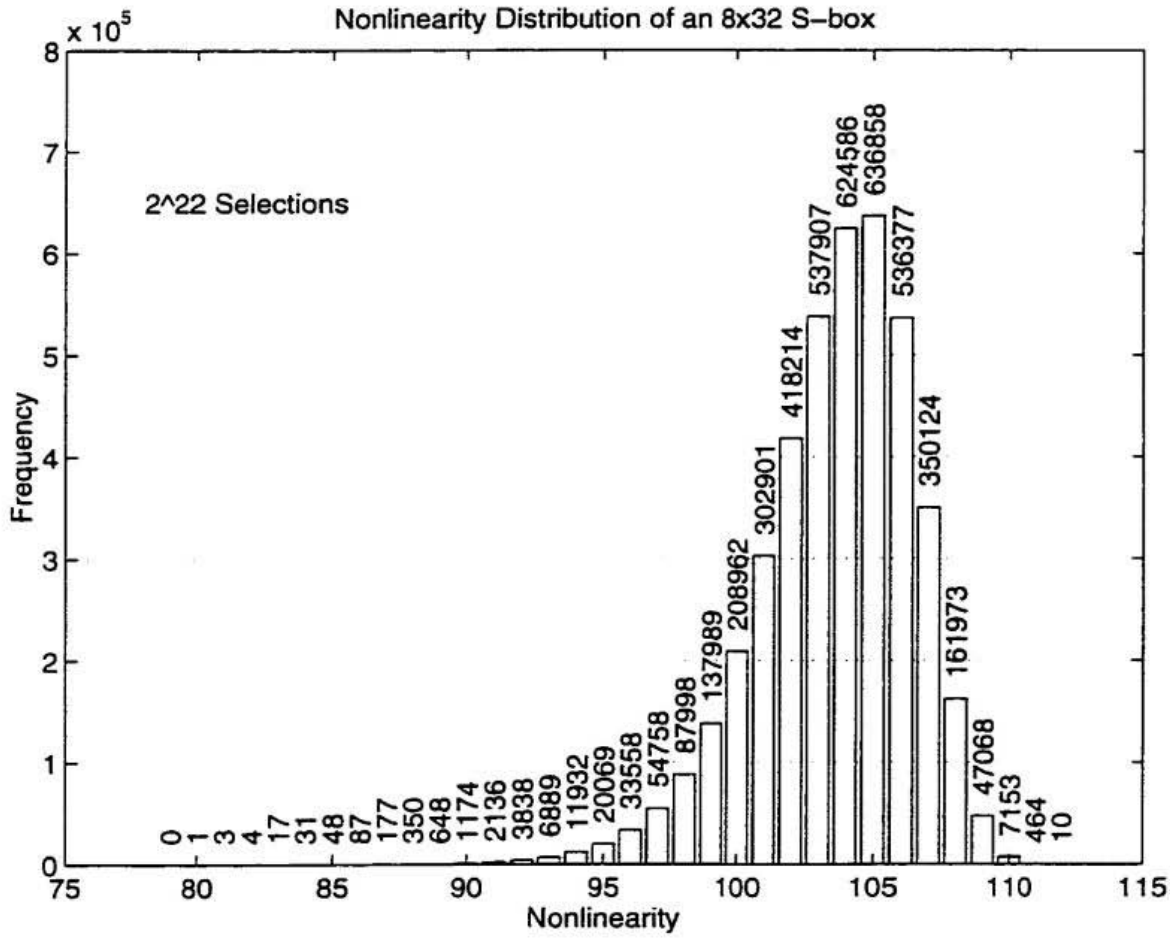

Figure 6.1: Nonlinearity Distribution of an 8×32 S-box with 2^{22} Selections

some probability distribution.

Assume that there are c categories ($c > 1$) and the expected frequency in each of these categories is denoted as $E_1, E_2, \ldots, E_c$. Similarly, the c observed frequencies are denoted as $O_1, O_2, \ldots, O_c$. To test the goodness of fit of the observed frequencies, O_i, to the expected frequencies, E_i, we use

$$\chi^2_{(c-1)} = \sum_{i=1}^{c} \frac{(O_i - E_i)^2}{E_i} \tag{6.5}$$

where $\chi^2_{(c-1)}$ is a chi-square variable with $(c-1)$ degrees of freedom. When O_i and

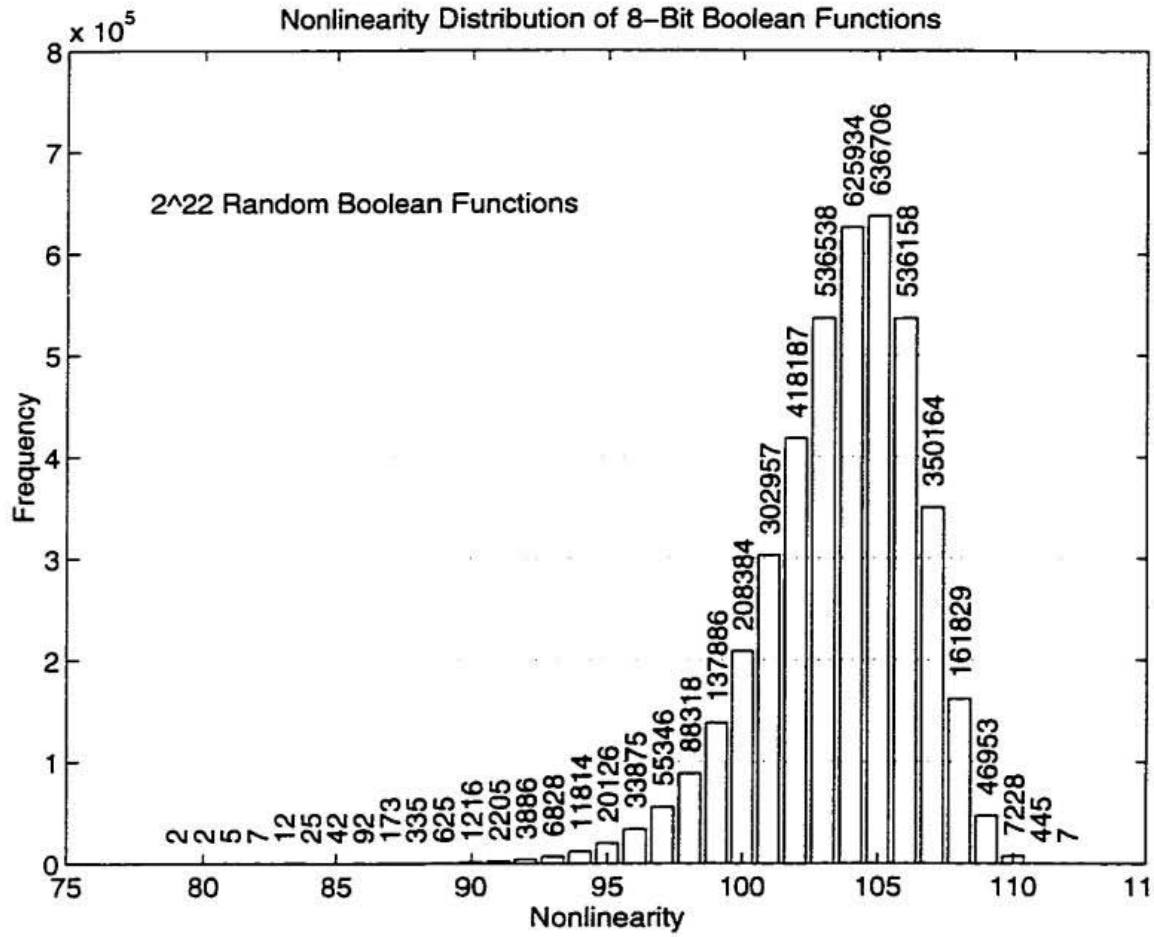

Figure 6.2: Nonlinearity Distribution of 2^{22} Boolean Functions

E_i are generally close, the value of χ^2 will be small. Conversely, if O_i and E_i are not close, the value of χ^2 will be large. Thus, the critical region for the statistical test given by Equation 6.5 will always be in the upper tail of the χ^2 distribution.

Consider two sets of nonlinearity distributions of Figure 6.1 and Figure 6.2. We assume that the values of the observed frequencies, O_i, are from Figure 6.1, and the values of the expected frequencies, E_i, are from Figure 6.2. Then $\chi^2 = 39$ with 33 degrees of freedom. By looking up the χ^2 distribution table [9], we find that $P(\chi^2_{33} \geq 47.1) = 0.05$. Thus, at the 0.95 level of significance, we can say that both

nonlinearity distributions are very close, based on the sampling data we have. This result further supports the assumption that the linearity of each boolean function derived by combining output bit boolean functions of an S-box can be based on considering each boolean function as independent.

6.2 Linear Cryptanalysis of the Ciphers

The objective of linear cryptanalysis is to find a linear approximation of a cipher only derived from plaintext, ciphertext, and key terms. A general linear approximation of a cipher is derived by combining a number of linear approximations of the S-boxes of different rounds so that intermediate terms are canceled. These linear approximations of the S-boxes usually contain both input and output bits. For a round function constructed from small S-boxes, such as in DES, it is feasible to find a best linear approximation of the cipher by completely searching all possible combinations of the linear approximations of the S-boxes of different rounds. However, this method is difficult to apply to the round function constructed with larger S-boxes, such as in the original CAST cipher.

In this section, we first introduce a linear approximation structure called an iterative linear approximation for a round function, and apply this structure to attack our unbalanced CAST ciphers with typical parameters.

6.2.1 Iterative Linear Approximation

The S-boxes in the round function of DES are implemented in parallel; their output bits are directed to the output bits of the round function with only a change of their

position because of the permutation. The nonlinearity of each S-box can be directly utilized by linear cryptanalysis to construct the linear approximation of the round function. Therefore, the nonlinearity of the S-box is much more important for the cipher to resist linear cryptanalysis. This is also true for the case of unbalanced CAST ciphers with $M = 1$.

On the other hand, for the original CAST or unbalanced CAST ciphers with $M > 1$, the output bit of the round function is XORed with the corresponding output bits of all S-boxes. Therefore, the nonlinearity of the round function is the result of the nonlinearities of all S-boxes in the round function. Even if one or more S-boxes have low nonlinearities, the combination of all S-boxes may still have high nonlinearity. This feature makes the round function of the original CAST stronger than the one of DES.

In Section 6.1.1, we have analyzed the nonlinearity distributions of the S-boxes for nonlinearities less than 2^{m-2}. We still do not know the complete nonlinearity distributions of the S-boxes because the exhaustive calculation is time-consuming work. Assuming a lower bound of the nonlinearity of 64 for $8 \times n$ S-boxes having $n \leq 32$ appears to be acceptably high and is not known to be tight. However, even if there exist nonlinearities less than 64 for $8 \times n$ S-boxes having $n \geq 32$, we still do not know whether these nonlinearities are useful to construct the best linear approximations of the round function and the cipher.

We consider an interesting linear approximation structure called an N/l-round iterative linear approximation, which has the form that the value of the XOR sum

of a subset of output bits of a round function is equal to 0 or 1 with a probability significantly different from 1/2. This form is shown as

$$\bigoplus_{j=1}^{a} Z_{i_j} = b \tag{6.6}$$

where $b \in \{0, 1\}$, Z_{i_j} represents i_j-th output bit of the round function, and a represents the number of output bits involved in the iterative linear approximation. Note that an iterative linear approximation does not involve any input bits of the round function. As a result, concatenating an iterative linear approximation to itself any number of times does not introduce any intermediate terms, and has a fixed reduction rate of the probability for each additional iterative linear approximation. In fact, an iterative linear approximation is quite similar to an iterative characteristic of differential cryptanalysis at a structural level. Figure 6.3 shows such a 2-round iterative linear approximation applicable to balanced CAST ciphers. In Figure 6.3, $\Gamma = R_{i_1} \oplus \ldots \oplus R_{i_a}$, where R_{i_j} represents i_j-th bit of the right half block of the input to a cipher round, and $\Phi = Z_{i_1} \oplus \ldots \oplus Z_{i_a}$, where $i_1, \ldots, i_a$ denote fixed bit positions of the right half block, R, and the output of the round function, Z.

If an unbalanced CAST cipher has $M = 1$, the output bits of the round function are directly the output bits of the S-box. Then the probability that the XOR sum of a subset of output bits of the round function is equal to 0 or 1 can be derived by calculating the hamming weight of the XOR sum of the corresponding subset of m-bit boolean functions of the S-box. If an unbalanced CAST cipher has $M > 1$, the output bits of the round function are derived from the XOR sum of the corresponding output bits of all S-boxes. Then the probability that the XOR sum of a subset of

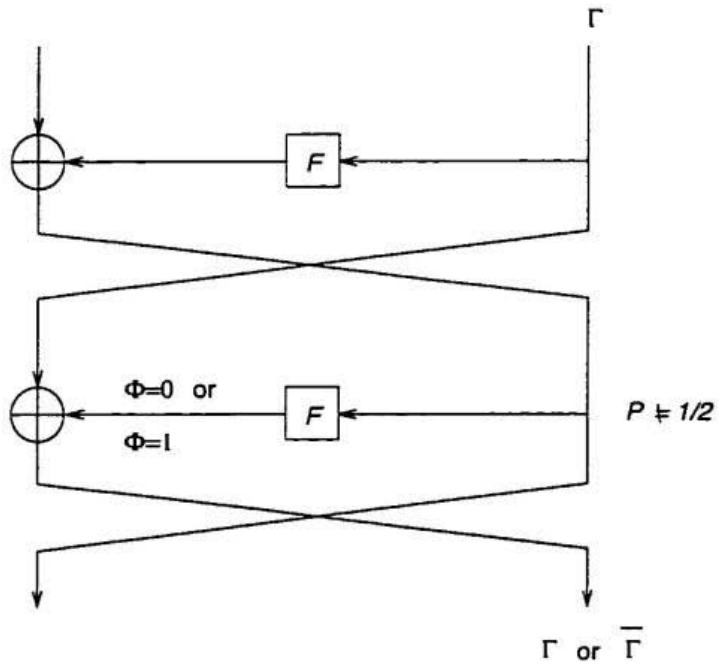

Figure 6.3: 2-Round Iterative Linear Approximation

output bits of the round function is equal to 0 or 1 can be determined by calculating the hamming weight of the XOR sum of the corresponding subset of m-bit boolean functions of every S-box and combining them with Matsui's Piling-up Lemma [19]. An unbalanced CAST cipher can have many iterative linear approximations of the form of Equation 6.6. The best one has the probability farthest from 1/2. Matsui's Piling-up Lemma is described as the following:

Lemma 6.1 (Piling-up Lemma) *Let X_i $(1 \leq i \leq n)$ be independent random variables whose values are 0 with probability p_i or 1 with probability $1 - p_i$. Then the probability that $X_1 \oplus X_2 \oplus \ldots \oplus X_n = 0$ is*

$$1/2 + 2^{n-1} \prod_{i=1}^{n} (p_i - 1/2).$$

Let f_i be an m-bit boolean function generated by XORing a subset of n m-bit boolean functions of output bits of an $m \times n$ S-box S_i and w_i be the hamming weight of f_i, where $i = 1, \cdots, M$. For a randomly selected $X \in \{0,1\}^m$, the probability of $f_i = 0$ is given by

$$P(f_i = 0) = 1 - \frac{w_i}{2^m}, \tag{6.7}$$

or the probability of $f_i = 1$ is given by

$$P(f_i = 1) = \frac{w_i}{2^m}. \tag{6.8}$$

Since each S-box is independently and randomly selected, using Matsui's Piling-up Lemma, we have a linear approximation for the round function

$$\bigoplus_{i=1}^{M} f_i = 0 \tag{6.9}$$

which holds with a probability

$$p_e = \frac{1}{2} + 2^{M-1} \prod_{i=1}^{M} (\frac{1}{2} - \frac{w_i}{2^m}) \tag{6.10}$$

where all f_i must involve the same subset of output bits of the S-boxes.

In general, an unbalanced CAST cipher should have a number of rounds, R, equal to an integer multiple of N/l. Assuming $R = r \cdot N/l$, then the structure of Equation 6.9 will be used n_f times by linear cryptanalysis to form an N/l-round iterative linear approximation, where $n_f = 1/2 \cdot N/l \cdot h/32$ for $h \leq 32$. Assuming that the probability of Equation 6.9 is bounded by $|p_e - 1/2| \leq \rho$, using Matsui's Piling-up Lemma, then the probability of an N/l-round iterative linear approximation, P_E, is bounded by

$$|P_E - 1/2| \leq 2^{n_f - 1} \cdot \rho^{n_f}. \tag{6.11}$$

Finally, an R-round unbalanced CAST cipher will have

$$|P_L - 1/2| \leq 2^{r-1} \cdot |P_E - 1/2|^r \tag{6.12}$$

where P_L is defined in Section 3.2, and the number of plaintexts required in the attack will be

$$N_L \geq 2^{2-2r} |P_E - 1/2|^{-2r} \tag{6.13}$$

determined as in Section 3.2. It is obvious that if one of f_i has the hamming weight close to 2^{m-1}, p_e will become close to $1/2$, and N_L will increase drastically.

6.2.2 Application of the Attack to Specific Ciphers

As we discuss in Section 6.1.2, f can be seen as a randomly generated m-bit boolean function. The probability that f has the hamming weight of w is given by

$$P(wt(f) = w) = \binom{2^m}{w} / 2^{2^m}. \tag{6.14}$$

and the probability that $wt(f) < w$ is given by

$$P(wt(f) < w) = \sum_{j=0}^{j=w-1} \binom{2^m}{j} / 2^{2^m}. \tag{6.15}$$

The probability distribution of w is a binomial distribution with mean $\mu = 2^{m-1}$ and variance $\sigma = 2^{(m-2)/2}$, and can be approximated by a normal distribution with the same mean and variance. According to our analysis in Section 6.1.2, there are 2^n randomly and independently generated m-bit boolean functions for an $m \times n$ S-box. Therefore, the probability that an $m \times n$ S-box has at least one m-bit boolean function whose hamming weight is less than w or greater than $2^m - w$ is given by

$$p = 1 - (1 - 2 \cdot \theta)^{2^n} \tag{6.16}$$

S-box $m \times n$	Hamming Weight w	Probability $P(wt(f) < w)$	Probability $1 - p$
8×16	84	9.59×10^{-9}	99.87%
8×32	72	3.39×10^{-13}	99.71%
8×56	56	5.04×10^{-21}	99.93%

Table 6.2: S-boxes with All Boolean Functions Having Hamming Weights Greater Than w and Less Than $256 - w$ for a Certain Probability

where $\theta = P(wt(f) < w)$ and $w \leq 2^{m-1}$.

First, we consider an 8×16 S-box. Assuming that $w = 84$, then $\theta = 9.59 \times 10^{-9}$ and $p = 1.3 \times 10^{-3}$ by calculating Equation 6.15 and Equation 6.16. Hence, we expect that all 8-bit boolean functions of the 8×16 S-box have the hamming weight greater than 83 and less than 173 with a probability of 99.87%.

Secondly, for an 8×32 S-box, assuming that $w = 72$, we have $\theta = 3.39 \times 10^{-13}$ and $p = 2.9 \times 10^{-3}$. Therefore, all 8-bit boolean functions of the 8×32 S-box are expected to have the hamming weight greater than 71 and less than 185 with a probability of 99.71%.

Finally, assuming that an 8×56 S-box has $w = 56$, then $\theta = 5.04 \times 10^{-21}$ and $p = 7.3 \times 10^{-4}$. As a result, we expect that all 8-bit boolean functions of the 8×56 S-box have the hamming weight greater than 55 and less than 201 with a probability of 99.93%.

A summary for $8 \times n$ S-boxes with all 8-bit boolean functions having the hamming weights greater than w and less than $256 - w$ for a certain probability is given in Table 6.2. It can be seen that a large S-box may have a boolean function whose hamming weight is farther from 2^{m-1} than a small S-box with a similar probability. This boolean function may be used to construct a useful linear approximation.

For an unbalanced CAST cipher with $M = 1$, $m = 8$, and $n = 16$, substituting $w = 84$ into Equation 6.10, we have $|p_e - 1/2| \leq \rho = 2^{-2.5}$. Since $n_f = 2$ for each eight rounds, from Equation 6.11, the probability of an 8-round iterative linear approximation is bounded by $|P_E - 1/2| \leq 2^{-4}$. As a result, the number of plaintexts required for linear cryptanalysis to determine one equivalent key bit with a 97.7% confidence level is at least 2^{62} for an 80-round cipher, and 2^{68} for an 88-round cipher.

If an unbalanced CAST cipher has $M = 2$, $m = 8$, and $n = 16$ and assuming that all 8-bit boolean functions of two 8×16 S-boxes are bounded by $w = 84$, then we have $|p_e - 1/2| \leq \rho = 2^{-4}$, which is the same as the probability of a 4-round iterative linear approximation since $n_f = 1$ for each four rounds. The number of plaintexts required to determine one equivalent key bit is at least 2^{62} for a 40-round cipher, and 2^{68} for a 44-round cipher.

Similarly, for $w = 72$ for an unbalanced CAST cipher with $M = 1$, $m = 8$, and $n = 32$, we have $|p_e - 1/2| \leq \rho = 2^{-2.2}$. Since $n_f = 4$ for each eight rounds, the probability of an 8-round iterative linear approximation is bounded by $|P_E - 1/2| \leq 2^{-5.8}$. The number of plaintexts required to determine one equivalent key bit is at least 2^{60} for a 48-round cipher, and 2^{69} for a 56-round cipher.

Also, for an unbalanced CAST cipher with $M = 2$, $m = 8$, and $n = 32$, we assume that $w = 72$, resulting in $|p_e - 1/2| \leq \rho = 2^{-3.4}$. Since $n_f = 2$ for each four rounds, the probability of a 4-round iterative linear approximation is bounded by $|P_E - 1/2| \leq 2^{-5.8}$. Clearly, the number of plaintexts required to determine one equivalent key bit is at least 2^{60} for a 24-round cipher, and 2^{69} for a 28-round cipher.

$m \times n$	M	$\lvert p_e - 1/2 \rvert$	N/l	n_f	$\lvert P_E - 1/2 \rvert$
8×16	1	$2^{-2.5}$	8	2	2^{-4}
	2	2^{-4}	4	1	2^{-4}
8×32	1	$2^{-2.2}$	8	4	$2^{-5.8}$
	2	$2^{-3.4}$	4	2	$2^{-5.8}$
	4	$2^{-5.8}$	2	1	$2^{-5.8}$
8×56	1	$2^{-1.8}$	8	7	$2^{-6.6}$

Table 6.3: Summary of the Probabilities of N/l-Round Iterative Linear Approximations

In the same way, for a balanced CAST cipher with $M = 4$, $m = 8$, and $n = 32$, assuming that $w = 72$ for all four 8×32 S-boxes, then we have $|p_e - 1/2| \leq \rho = 2^{-5.8}$, which is the same as the probability of a 2-round iterative linear approximation since $n_f = 1$ for each two rounds. The number of plaintexts required to determine one equivalent key bit is at least 2^{60} for a 12-round cipher, and 2^{69} for a 14-round cipher. In comparison, the loose analysis of [14] implied that a 12-round original CAST cipher needs at least 2^{50} known plaintexts.

If an unbalanced CAST cipher has $M = 1$, $m = 8$, and $n = 56$, substituting $w = 56$ into Equation 6.10, we have $|p_e - 1/2| \leq \rho = 2^{-1.8}$. The probability of an 8-round iterative linear approximation is bounded by $|P_E - 1/2| \leq 2^{-6.6}$ since $n_f = 7$ for each eight rounds. The number of plaintexts required to determine one equivalent key bit is at least 2^{58} for a 40-round cipher, and 2^{69} for a 48-round cipher.

A summary of the probabilities of N/l-round iterative linear approximations is given in Table 6.3. Note that all N/l-round iterative linear approximations have the same probability for unbalanced CAST ciphers whose round functions are constructed by the same type of S-box. Although an 8×16 S-box has a larger w and smaller $|p_e - 1/2|$ than an 8×56 S-box for the similar probability given in Table 6.2, the N/l-

round iterative linear approximation made of 8×16 S-boxes has a higher probability than the one made of 8×56 S-boxes. This is because only two bytes of the left half block are XORed by the output bits of the round function constructed by 8×16 S-boxes in each round, while seven bytes of the data are XORed by the output bits of the round function constructed by 8×56 S-boxes in each round.

Note that Table 6.3 lists the upper bounds of the probabilities of N/l-round iterative linear approximations. The real probabilities will be much smaller in a practical linear cryptanalysis. This is because the rotation operation causes the plaintext bit positions to be changed every round for an unbalanced CAST cipher and an N/l-round iterative linear approximation is combined by n_f different subsets of output bits of the round function. If one of the output bit boolean functions has the hamming weight close to $1/2$, the probability of the N/l-round iterative linear approximation will become close to $1/2$ rapidly, and the iterative linear approximation will be useless for linear cryptanalysis. Therefore, large n_f or M help the unbalanced CAST ciphers become resistant to linear cryptanalysis.

A summary of linear cryptanalysis based on iterative linear approximations given in Table 6.3 to unbalanced CAST ciphers with different parameters is shown in Table 6.4. Interestingly, if two ciphers satisfy Equation 4.2, they will have the same security level according to the analysis techniques used here. It should be noted that the number of known plaintexts required to linearly cryptanalyze a cipher can not exceed 2^{64} if the block size of the cipher is equal to 64 bits. For the cases in the table where the plaintexts needed exceed 2^{64}, the implication is that such a cipher is

S-box		Number of	Plaintexts
$m \times n$	M	Rounds	Needed
8×16	1	80	2^{62}
		88	2^{68}
	2	40	2^{62}
		44	2^{68}
8×32	1	48	2^{60}
		56	2^{69}
	2	24	2^{60}
		28	2^{69}
	4	12	2^{60}
		14	2^{69}
8×56	1	40	2^{58}
		48	2^{69}

Table 6.4: Summary of Linear Cryptanalysis with Iterative Linear Approximations

theoretically secure against linear cryptanalysis.

6.3 Conclusion

We have considered the linear cryptanalysis of unbalanced CAST ciphers. First of all, we have analyzed the nonlinearity of S-boxes which are used in our unbalanced CAST ciphers. It can be seen that a large S-box has a lower nonlinearity than a small S-box and there is a trade-off between the nonlinearity requirement and the XOR distribution requirement. Then, we have provided an experimental justification for the assumption that 2^n m-bit boolean functions linearly combined by n m-bit boolean functions are randomly selected. Finally, we have introduced the concept of the iterative linear approximation and applied it to the attack of unbalanced CAST ciphers. The analysis result shows that a 48-round unbalanced CAST cipher with one 8×32 S-box and a 24-round unbalanced CAST cipher with two 8×32 S-box are secure against linear cryptanalysis.

Chapter 7

Information Theoretic View of Unbalanced CAST Ciphers

In this chapter, we present an analysis of unbalanced CAST ciphers from an information theoretic point of view. First, we derive simple equations of calculating an entropy and information leakage. Second, we apply these equations to analyze the information leakages on the round function level for unbalanced CAST ciphers. Finally, we analyze the information leakages on the cipher level.

7.1 Information Leakage

Entropy and information leakage are defined in Section 2.3.2 by Equation 2.8 to Equation 2.11, and can be used as one measure of the cryptographic strength of a cipher. In this section, we discuss entropy and information leakage in detail.

Define a random variable Z, such that $Z \in \{0, 1\}$. By Equation 2.8, the entropy of Z is given by

$$H(Z) = -[P(Z = 0) \log P(Z = 0) + P(Z = 1) \log P(Z = 1)] \qquad (7.1)$$

where $P(Z = 0)$ is the probability of $Z = 0$, $P(Z = 1)$ is the probability of $Z = 1$,

and the base of the log is 2. Letting $|P(Z=0)-\frac{1}{2}| = \varepsilon_z$, then

$$H(Z) = -[(\frac{1}{2}+\varepsilon_z)\log(\frac{1}{2}+\varepsilon_z) + (\frac{1}{2}-\varepsilon_z)\log(\frac{1}{2}-\varepsilon_z)]. \tag{7.2}$$

Using a Taylor's series expansion,

$$\log(\frac{1}{2}+\varepsilon) = -1 + 2\varepsilon - 2\varepsilon^2 + \frac{8}{3}\varepsilon^3 - \cdots. \tag{7.3}$$

Substituting Equation 7.3 into Equation 7.2, and assuming that ε_z is sufficiently small, we have

$$H(Z) \approx 1 - 2\varepsilon_z^2. \tag{7.4}$$

It is obvious that if $\varepsilon_z \to 0$, $H(Z) \to 1$, which means that the random variable, Z, has maximum uncertainty of its outcome.

Define another random variable $X \in \{0,1\}$ such that $P(X=0) = P(X=1) = \frac{1}{2}$. By Equation 2.9, the conditional entropy of Z given X is

$$H(Z|X) = -\sum_{X=0}^{1} P(X)\left[\sum_{Z=0}^{1} P(Z|X)\log P(Z|X)\right]. \tag{7.5}$$

Let $|P(Z=0|X=0)-\frac{1}{2}| = \varepsilon_{z|0}$ and $|P(Z=0|X=1)-\frac{1}{2}| = \varepsilon_{z|1}$, and assume that $\varepsilon_{z|0}$ and $\varepsilon_{z|1}$ are sufficiently small. Using Equation 7.3 again, we have

$$H(Z|X) \approx 1 - \varepsilon_{z|0}^2 - \varepsilon_{z|1}^2. \tag{7.6}$$

The information leakage or mutual information between random variables X and Z is defined as

$$I(Z,X) = H(Z) - H(Z|X). \tag{7.7}$$

An information leakage equal to zero implies that information about X does not reduce the uncertainty of Z, and vice versa. In an ideal cipher, we are interested in

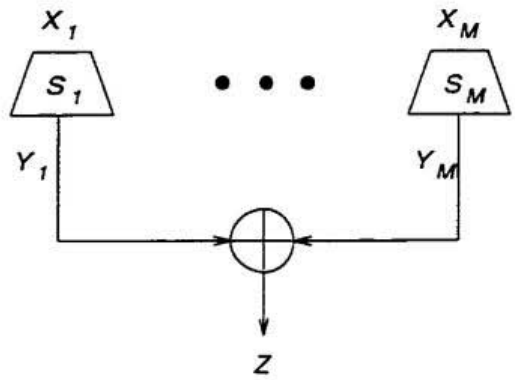

Figure 7.1: Bit Structure of the Round Function

$H(Z) = 1$ and $I(Z, X) = 0$, where Z represents a ciphertext bit and X represents a plaintext bit.

7.2 Information Leakages of Round Functions

An unbalanced CAST cipher employs M $m \times n$ S-boxes in its round function. Each output bit of the round function is generated by XORing corresponding output bits of all S-boxes. Figure 7.1 shows the bit structure of the round function of an unbalanced CAST cipher, where X_i represents one of input bits of an $m \times n$ S-box S_i, Y_i represents one of output bits of the S-box, and Z represents one of output bits of the round function, where $Z = \oplus_{i=1}^{M} Y_i$.

An $m \times n$ S-box may be viewed as a $2^m \times n$ binary matrix, in which each column represents an output bit for the inputs corresponding to each row. Since each S-box is randomly and independently generated, the probability that each bit is zero or one in the binary matrix is $1/2$. Therefore, the number of zeros for each output bit in the binary matrix is a binomial distribution with $G_1 = 2^m$ and $p_1 = q_1 = 1/2$, whose mean is $\mu_1 = G_1 p_1 = 2^{m-1}$ and variance is $\sigma_1 = \sqrt{G_1 p_1 q_1} = 2^{(m-2)/2}$.

Assume that $|P(Y_i = 0) - 1/2|$ is bounded by ε_1 with a confidence level of ν_1. Let $|P(Z = 0) - 1/2| = \varepsilon_z$. Using Matsui's Piling-up Lemma [19], we have

$$\varepsilon_z \leq 2^{M-1}\varepsilon_1^M. \tag{7.8}$$

Substituting Equation 7.8 into Equation 7.4, we have the unconditional entropy of Z

$$H(Z) \geq 1 - 2^{2M-1}\varepsilon_1^{2M}. \tag{7.9}$$

Similarly, in the binary matrix, there are 2^{m-1} rows corresponding to a value of zero and 2^{m-1} rows corresponding to a value of one for a particular input bit. Given an input bit equal to zero or one, the number of zeros for an output bit is a binomial distribution with $G_2 = 2^{m-1}$ and $p_2 = q_2 = 1/2$, whose mean is $\mu_2 = G_2 p_2 = 2^{m-2}$ and variance is $\sigma_2 = \sqrt{G_2 p_2 q_2} = 2^{(m-3)/2}$.

Assume that $|P(Y_i = 0|X_i = 0) - 1/2|$ and $|P(Y_i = 0|X_i = 1) - 1/2|$ are bounded by ε_2 with a confidence level of ν_2. Let $|P(Z = 0|X_i = 0) - 1/2|$ be $\varepsilon_{z|0}$, and $|P(Z = 0|X_i = 1) - 1/2|$ be $\varepsilon_{z|1}$. Since M $m \times n$ S-boxes are randomly and independently generated, $P(Y_j|X_i) = P(Y_j)$, where $j \neq i$. Assume that $|P(Y_j = 0) - 1/2|$ is bounded by ε_1 with a confidence level of ν_1. Using Matsui's Piling-up Lemma, we have

$$\varepsilon_{z|0} = \varepsilon_{z|1} \leq 2^{M-1}\varepsilon_1^{M-1}\varepsilon_2. \tag{7.10}$$

Substituting Equation 7.10 into Equation 7.6, we have the conditional entropy of Z given X_i

$$H(Z|X_i) \geq 1 - 2^{2M-1}\varepsilon_1^{2M-2}\varepsilon_2^2. \tag{7.11}$$

Since all entropies are less than or equal to one and $H(Z) \geq H(Z|X_i)$, by Equation 7.7, the static input-output bit information leakage between Z and X_i is bounded

by

$$0 \leq I(Z; X_i) \leq 2^{2M-1}\varepsilon_1^{2M-2}\varepsilon_2^2. \tag{7.12}$$

It can be seen that the magnitude of the information leakage between an input bit and an output bit of a round function mainly depends on the number of input bits of an S-box, m, which influences ε_i, and the number of the S-boxes, M, which the round function employs.

Consider now the computation of the confidence level with which Equation 7.9 and Equation 7.11 hold. Assume that a random variable z represents the number of zeros in a randomly generated G-bit binary vector Y, and satisfies a binomial distribution with mean μ and variance σ. Let $P(z \leq z_0) = \alpha$, where $z_0 \leq \mu$. Then $P(|z - \mu| < \mu - z_0) = 1 - 2\alpha$, which means that $|z - \mu|$ is less than $\mu - z_0$ with a confidence level of $1 - 2\alpha$. Since $z/G = z/(2\mu)$ represents the probability of $Y = 0$, we may say that $|P(Y = 0) - 1/2|$ is less than ε with a confidence level of ν, where $\varepsilon = 1/2 - z_0/G$ and $\nu = 1 - 2\alpha$. Meanwhile, if there are n such randomly generated G-bit binary vectors, the confidence level, or the probability, with which all Ys satisfy $|P(Y = 0) - 1/2| < \varepsilon$, will be $(1 - 2\alpha)^n \approx 1 - 2n\alpha$ if $2n\alpha \ll 1$.

Consider an 8×56 S-box, the number of zeros for each output bit is a binomial distribution with $G_1 = 2^8$ and $p_1 = q_1 = 1/2$, whose mean $\mu_1 = 2^7$ and variance $\sigma_1 = 2^3$. For $z_0 = 97$, we have $\alpha = 6.4 \times 10^{-5}$, and hence $\varepsilon_1 = 1/2 - 98/256 \approx 2^{-3}$ with a confidence level of $\nu_1 = 1 - 2 \cdot 56 \cdot 6.4 \times 10^{-5} \approx 99.3\%$. Therefore, we may say that $|P(Y_i = 0) - 1/2|$ is bounded by $\varepsilon_1 = 2^{-3}$ with a confidence level of $\nu_1 = 99.3\%$ for an 8×56 S-box.

In the same way, given an input bit equal to zero or one, the number of zeros for an output bit is a binomial distribution with $G_2 = 2^7$ and $p_2 = q_2 = 1/2$, whose mean $\mu_2 = 2^6$ and variance $\sigma_2 = 2^{2.5}$. For $z_0 = 39$, we have $\alpha = 5.8 \times 10^{-6}$. Then $\varepsilon_2 = 1/2 - 40/128 \approx 2^{-2.4}$ with a confidence level of $\nu_2 = 1 - 2 \cdot 2 \cdot 8 \cdot 56 \cdot 5.8 \times 10^{-6} \approx 99.0\%$. Therefore, we may say that $|P(Y_i = 0|X_i = 0) - 1/2|$ and $|P(Y_i = 0|X_i = 1) - 1/2|$ are bounded by $\varepsilon_2 = 2^{-2.4}$ with a confidence level of $\nu_2 = 99.0\%$ for an 8×56 S-box.

Since the sizes of all S-boxes are smaller than 8×56 in our analysis and corresponding confidence levels will be higher for the selected values of ε_1 and ε_2, we assume that ε_1, ν_1, ε_2, and ν_2 for an 8×56 S-box are suitable for all remaining $8 \times n$ S-boxes. From Equation 7.9, Equation 7.11, and Equation 7.12, the unbalanced CAST ciphers with one 8×16, one 8×32, and one 8×56 S-box all have $H(Z) \geq 1 - 2^{-5}$, $H(Z|X_i) \geq 1 - 2^{-3.8}$, and $I(Z, X_i) \leq 2^{-3.8}$ for their round functions with a high level of confidence. The unbalanced CAST ciphers with two 8×16 and two 8×32 S-boxes both have $H(Z) \geq 1 - 2^{-9}$, $H(Z|X_i) \geq 1 - 2^{-7.8}$, and $I(Z, X_i) \leq 2^{-7.8}$ for their round functions. A balanced CAST cipher with four 8×32 S-boxes has $H(Z) \geq 1 - 2^{-17}$, $H(Z|X_i) \geq 1 - 2^{-15.8}$, and $I(Z, X_i) \leq 2^{-15.8}$ for its round function. It is obvious that a round function constructed by more S-boxes has less information leakage.

The above analysis has assumed that all input bits of the round functions are randomly and independently changed. We believe that the assumption is reasonable. As we have discussed in Section 4.3, each output bit of the round function of an unbalanced CAST cipher is influenced by all the input bits and the rotation operation immediately brings some or all output bits of the round function to the input position

$m \times n$	M	$H(Z)$	$H(Z\|X_i)$	$I(Z, X_i)$
8×16	1	$1 - 2^{-5}$	$1 - 2^{-3.8}$	$2^{-3.8}$
	2	$1 - 2^{-9}$	$1 - 2^{-7.8}$	$2^{-7.8}$
8×32	1	$1 - 2^{-5}$	$1 - 2^{-3.8}$	$2^{-3.8}$
	2	$1 - 2^{-9}$	$1 - 2^{-7.8}$	$2^{-7.8}$
	4	$1 - 2^{-17}$	$1 - 2^{-15.8}$	$2^{-15.8}$
8×56	1	$1 - 2^{-5}$	$1 - 2^{-3.8}$	$2^{-3.8}$

Table 7.1: Summary of Static Input-Output Bit Information Leakages of Round Functions of Unbalanced CAST Ciphers

of the round function in the next round. Therefore, after two rounds all input bits of the round function will be changed.

A summary of static input-output bit information leakages of round functions is given in Table 7.1.

As a comparison, we investigate the static input-output bit information leakage of the round function of DES. Since the output bits of the round function are directly the output bits of the S-boxes, information leakages between input and output bits of the round function of DES can be obtained by examining each S-box directly. Although all output bits of the eight 6×4 S-boxes in DES are zero balanced, which means that the unconditional entropy of each output bit of the round function is exactly one, its conditional entropy is not. Table 7.2 can be used to derive the conditional probabilities of the S-box, $S1$, where x_i represents the i-th input bit, y_j represents the j-th output bit, p_0 is used to determine $P(y_j = 0|x_i = 0)$, and p_1 is used to determine $P(y_j = 0|x_i = 1)$. The real conditional probability in Table 7.2 is calculated by $P(y_j|x_i) = 1/2 + entry/32$. For example, $P(y_2 = 0|x_3 = 0) = 1/2 - 3/32$.

By looking at all eight S-boxes, we find that $|P(y_j = 0|x_i = 0) - 1/2| = |P(y_j = 0|x_i = 1) - 1/2| = 3/32 \approx 2^{-3.4}$ is a maximum value. Therefore, by Equation 7.5

	y_1		y_2		y_3		y_4	
	p_0	p_1	p_0	p_1	p_0	p_1	p_0	p_1
x_1	0	0	0	0	0	0	0	0
x_2	0	0	-1	1	1	-1	1	-1
x_3	-2	2	-3	3	-1	1	2	-2
x_4	-1	1	-1	1	-1	1	1	-1
x_5	1	-1	-1	1	-1	1	-1	1
x_6	0	0	0	0	0	0	0	0

Table 7.2: Conditional Probabilities of $S1$ in the Round Function of DES

and Equation 7.7, the maximum static input-output bit information leakage of the round function of DES is $2.55 \times 10^{-2} \approx 2^{-5.3}$, which is much greater than the balanced CAST, almost equivalent to unbalanced CAST ciphers with two 8×16 and two 8×32 S-boxes, and less than unbalanced CAST ciphers with one 8×16, one 8×32, and one 8×56 S-box.

Until now, we have discussed the static input-output bit information leakages of round functions of unbalanced CAST ciphers. The corresponding dynamic input-output bit information leakages, according to Figure 7.1, are defined as

$$I(\Delta Z, \Delta X_i) = H(\Delta Z) - H(\Delta Z | \Delta X_i) \tag{7.13}$$

where $\Delta X_i = X_i \oplus X_i^*$ is one of input bits of an S-box, $\Delta Y_i = Y_i \oplus Y_i^*$ is one of output bits of the S-box, and $\Delta Z = Z \oplus Z^*$ is one of output bits of the round function, where $\Delta Z = \oplus_{i=1}^{M} \Delta Y_i$. Since bit values of X_i and X_i^* can be selected independently, the bit values of Y_i and Y_i^*, Z and Z^* can be seen as independent. Define ε_z' as $|P(\Delta Z = 0) - \frac{1}{2}|$. Using Matsui's Piling-up Lemma, Equation 7.8 and Equation 7.9 can be rearranged to

$$\varepsilon_z' = 2\varepsilon_z^2 \leq 2^{2M-1}\varepsilon_1^{2M} \tag{7.14}$$

and

$$H(\Delta Z) \geq 1 - 2^{4M-1}\varepsilon_1^{4M}. \tag{7.15}$$

Similarly, define $\varepsilon'_{z|0}$ as $|P(\Delta Z = 0|\Delta X = 0) - \frac{1}{2}|$ and $\varepsilon'_{z|1}$ as $|P(\Delta Z = 0|\Delta X = 1) - \frac{1}{2}|$. Equation 7.10 and Equation 7.11 can be rearranged to

$$\varepsilon'_{z|0} = \varepsilon'_{z|1} \leq 2^{2M-1}\varepsilon_1^{2M-2}\varepsilon_2^2 \tag{7.16}$$

and

$$H(\Delta Z|\Delta X_i) \geq 1 - 2^{4M-1}\varepsilon_1^{4M-4}\varepsilon_2^4. \tag{7.17}$$

Therefore, the dynamic input-output bit information leakage is bounded by

$$0 \leq I(\Delta Z, \Delta X_i) \leq 2^{4M-1}\varepsilon_1^{4M-4}\varepsilon_2^4. \tag{7.18}$$

The unbalanced CAST ciphers with one 8×16, one 8×32, and one 8×56 S-box all have $H(\Delta Z) \geq 1 - 2^{-9}$, $H(\Delta Z|\Delta X_i) \geq 1 - 2^{-6.6}$, and $I(\Delta Z, \Delta X_i) \leq 2^{-6.6}$ for their round functions with a high level of confidence. The unbalanced CAST ciphers with two 8×16 and two 8×32 S-boxes both have $H(\Delta Z) \geq 1 - 2^{-17}$, $H(\Delta Z|\Delta X_i) \geq 1 - 2^{-14.6}$, and $I(\Delta Z, \Delta X_i) \leq 2^{-14.6}$ for their round functions with a high level of confidence. A balanced CAST cipher with four 8×32 S-boxes has $H(\Delta Z) \geq 1 - 2^{-33}$, $H(\Delta Z|\Delta X_i) \geq 1 - 2^{-30.6}$, and $I(\Delta Z, \Delta X_i) \leq 2^{-30.6}$ for its round function. A summary of dynamic input-output bit information leakages of round functions is given in Table 7.3.

As a comparison, the maximum dynamic input-output bit information leakage of the round function of DES is $8.92 \times 10^{-4} \approx 2^{-10}$.

$m \times n$	M	$H(\Delta Z)$	$H(\Delta Z \mid \Delta X_i)$	$I(\Delta Z, \Delta X_i)$
8×16	1	$1 - 2^{-9}$	$1 - 2^{-6.6}$	$2^{-6.6}$
	2	$1 - 2^{-17}$	$1 - 2^{-14.6}$	$2^{-14.6}$
8×32	1	$1 - 2^{-9}$	$1 - 2^{-6.6}$	$2^{-6.6}$
	2	$1 - 2^{-17}$	$1 - 2^{-14.6}$	$2^{-14.6}$
	4	$1 - 2^{-33}$	$1 - 2^{-30.6}$	$2^{-30.6}$
8×56	1	$1 - 2^{-9}$	$1 - 2^{-6.6}$	$2^{-6.6}$

Table 7.3: Summary of Dynamic Input-Output Bit Information Leakages of Round Functions of Unbalanced CAST Ciphers

7.3 Information Leakages of the Ciphers

In this section we extend the analysis of the information leakage from a one round level to a multiple round level. As previously, we assume that an unbalanced CAST cipher must have a number of rounds, R, equal to an integer multiple of N/l. Let this integer be r, so that $R = r \cdot N/l$. After every N/l rounds, each plaintext bit will be XORed with one of output bits of the round function $n_f = 1/2 \cdot N/l \cdot h/32$ times for $h \leq 32$. Each time the output bit of the round function may be different. For example, it can be seen by looking at Table 4.1 that the unbalanced CAST cipher with one 8×32 S-box has each plaintext bit XORed with one of output bits of its round function four times after every eight rounds.

In order to get worst-case upper bounds on the information leakage, let the left half block of plaintexts be unchanged and each bit of right half block of plaintexts be changed randomly, independently, and with a probability of 1/2. Assume that output blocks of the round function are independent between rounds. Using Matsui's Piling-up Lemma, from Equation 7.8 and Equation 7.9, an R-round unbalanced CAST

cipher with M $m \times n$ S-boxes will have

$$\varepsilon_z \leq 2^{Mn_f r-1} \varepsilon_1^{Mn_f r} \tag{7.19}$$

and

$$H(Z) \geq 1 - 2^{2Mn_f r-1} \varepsilon_1^{2Mn_f r} \tag{7.20}$$

where Z represents one bit of left half block of the ciphertext. Similarly, from Equation 7.10 and Equation 7.11, an R-round unbalanced CAST cipher with M $m' \times n$ S-boxes will have

$$\varepsilon_{z|0} = \varepsilon_{z|1} \leq 2^{Mn_f r-1} \varepsilon_1^{Mn_f r-1} \varepsilon_2 \tag{7.21}$$

and

$$H(Z|X) \geq 1 - 2^{2Mn_f r-1} \varepsilon_1^{2Mn_f r-2} \varepsilon_2^2 \tag{7.22}$$

where X represents one bit of right half block of the plaintext. Therefore, the static input-output bit information leakage between Z and X for an R-round unbalanced CAST cipher with M $m \times n$ S-boxes is bounded by

$$0 \leq I(Z, X) \leq 2^{2Mn_f r-1} \varepsilon_1^{2Mn_f r-2} \varepsilon_2^2. \tag{7.23}$$

As a result, a 64-round unbalanced CAST cipher with one 8×16 S-box, a 32-round unbalanced CAST cipher with two 8×16 S-boxes, a 32-round unbalanced CAST cipher with one 8×32 S-box, a 16-round unbalanced CAST cipher with two 8×32 S-boxes, and an 8-round balanced CAST cipher with four 8×32 S-boxes all have $H(Z) \geq 1 - 2^{-65}$, $H(Z|X) \geq 1 - 2^{-63.8}$, and $I(Z, X) \leq 2^{-63.8}$, since all of them have $Mn_f r = 16$.

$m \times n$	M	R	$H(Z)$	$H(Z\|X)$	$I(Z,X)$
8×16	1	64	$1-2^{-65}$	$1-2^{-63.8}$	$2^{-63.8}$
	2	32	$1-2^{-65}$	$1-2^{-63.8}$	$2^{-63.8}$
8×32	1	32	$1-2^{-65}$	$1-2^{-63.8}$	$2^{-63.8}$
	2	16	$1-2^{-65}$	$1-2^{-63.8}$	$2^{-63.8}$
	4	8	$1-2^{-65}$	$1-2^{-63.8}$	$2^{-63.8}$
8×56	1	24	$1-2^{-85}$	$1-2^{-83.8}$	$2^{-83.8}$

Table 7.4: Summary of Static Input-Output Bit Information Leakages for Multiple Round Unbalanced CAST Ciphers

The analysis of the unbalanced CAST cipher with one 8×56 S-box is a little different since there is no swapping operation in each round. However, the analysis can still be carried on by keeping one byte of plaintexts unchanged and letting all remaining plaintext bits change randomly, independently, and with a probability of $1/2$. Then, the condition that the output blocks of the round function are independent between rounds is reasonable. Since $n_f = 7$ for every eight rounds, a 24-round cipher has $Mn_f r = 21$, $H(Z) \geq 1-2^{-85}$, $H(Z|X) \geq 1-2^{-83.8}$, and $I(Z,X) \leq 2^{-83.8}$, which is better than an 32-round unbalanced CAST cipher with one 8×32 S-box.

A summary of static input-output bit information leakages for multiple round unbalanced CAST ciphers is given in Table 7.4.

The dynamic input-output bit information leakages for multiple round unbalanced CAST ciphers can be derived from the corresponding static information leakages in the same way as described in the previous section. From Equation 7.19 and Equation 7.20, an R-round unbalanced CAST cipher with M $m \times n$ S-boxes will have

$$\varepsilon_z' \leq 2^{2Mn_f r-1} \varepsilon_1^{2Mn_f r} \tag{7.24}$$

and

$$H(\Delta Z) \geq 1 - 2^{4Mn_fr-1}\varepsilon_1^{4Mn_fr}. \tag{7.25}$$

Similarly, from Equation 7.21 and Equation 7.22, an R-round unbalanced CAST cipher with M $m \times n$ S-boxes will have

$$\varepsilon'_{z|0} = \varepsilon'_{z|1} \leq 2^{2Mn_fr-1}\varepsilon_1^{2Mn_fr-2}\varepsilon_2^2 \tag{7.26}$$

and

$$H(\Delta Z|\Delta X) \geq 1 - 2^{4Mn_fr-1}\varepsilon_1^{4Mn_fr-4}\varepsilon_2^4. \tag{7.27}$$

The corresponding dynamic input-output bit information leakage is bounded by

$$0 \leq I(\Delta Z, \Delta X) \leq 2^{4Mn_fr-1}\varepsilon_1^{4Mn_fr-4}\varepsilon_2^4. \tag{7.28}$$

As a result, a 64-round unbalanced CAST cipher with one 8×16 S-box, a 32-round unbalanced CAST cipher with two 8×16 S-boxes, a 32-round unbalanced CAST cipher with one 8×32 S-box, a 16-round unbalanced CAST cipher with two 8×32 S-boxes, and an 8-round balanced CAST cipher with four 8×32 S-boxes all have $H(\Delta Z) \geq 1 - 2^{-129}$, $H(\Delta Z|\Delta X) \geq 1 - 2^{-126.6}$, and $I(\Delta Z, \Delta X) \leq 2^{-126.6}$, since all of them have same $Mn_fr = 16$. An 24-round unbalanced CAST cipher with one 8×56 S-box has $Mn_fr = 21$, $H(\Delta Z) \geq 1 - 2^{-169}$, $H(\Delta Z|\Delta X) \geq 1 - 2^{-166.6}$, and $I(\Delta Z, \Delta X) \leq 2^{-166.6}$.

A summary of dynamic input-output bit information leakages for multiple round unbalanced CAST ciphers is given in Table 7.5.

Since every output bit of the round function of DES is zero balanced, its ε_1 is equal to zero and the unconditional entropy is equal to one. The static and dynamic

$m \times n$	M	R	$H(\Delta Z)$	$H(\Delta Z\|\Delta X)$	$I(\Delta Z, \Delta X)$
8×16	1	64	$1 - 2^{-129}$	$1 - 2^{-126.6}$	$2^{-126.6}$
	2	32	$1 - 2^{-129}$	$1 - 2^{-126.6}$	$2^{-126.6}$
8×32	1	32	$1 - 2^{-129}$	$1 - 2^{-126.6}$	$2^{-126.6}$
	2	16	$1 - 2^{-129}$	$1 - 2^{-126.6}$	$2^{-126.6}$
	4	8	$1 - 2^{-129}$	$1 - 2^{-126.6}$	$2^{-126.6}$
8×56	1	24	$1 - 2^{-169}$	$1 - 2^{-166.6}$	$2^{-166.6}$

Table 7.5: Summary of Dynamic Input-Output Bit Information Leakages for Multiple Round Unbalanced CAST Ciphers

input-output bit information leakages of DES will approach zero drastically after three rounds. It is possible to generate S-boxes for unbalanced CAST ciphers that have all their output bits to be zero balanced and hence make unbalanced CAST ciphers possess the same feature as DES. Furthermore, some unbalanced CAST ciphers with $M > 1$ will have zero information leakage for their round functions if such S-boxes are selected. For an S-box having eight input bits, the probability that a randomly generated 256-bit binary vector of an output bit of the S-box is zero balanced is $\binom{256}{128} / 2^{256} \approx 0.05$. Therefore, it is not difficult to generate an 8×32 S-box whose output bits are all zero balanced.

7.4 Conclusion

We have provided a method to compute the upper bounds of the static and dynamic input-output bit information leakages for both single round and multiple round unbalanced CAST ciphers. As a result, although a round function constructed by more S-boxes has a smaller information leakage, two ciphers will have the same information leakage if Equation 4.2 holds. For 2-round DES, the maximum static and dynamic input-output bit information leakages are $2^{-5.3}$ and 2^{-10}, respectively. On the other

hand, for N/l-round unbalanced CAST ciphers with M 8×32 S-boxes, the corresponding information leakages are bounded by $2^{-15.8}$ and $2^{-30.6}$, respectively. In order to further reduce the information leakages, we have recommended that all output bits of the S-boxes generated for unbalanced CAST ciphers should have the same number of zeros and ones.

It should be noted that we have only examined the information leakages at the single bit level. Since the unbalanced CAST ciphers employ a set of large S-boxes, the investigation of the information leakages on the multiple bit level becomes much more difficult. Furthermore, when analyzing the information leakages of the multiple round unbalanced CAST ciphers, we fix some plaintext bits to get the worst-case upper bounds of the information leakages. Otherwise, the analysis becomes trivial.

Chapter 8

Conclusions

DES is about twenty years old and is coming to the end of its useful life time. NIST is calling for a new standard to replace DES which offers a higher level of security and better efficiency. The original CAST cipher is one of a few proposed private-key block ciphers which has efficient and secure properties. In this thesis, we have presented a new class of unbalanced CAST ciphers which requires less memory than the original CAST cipher and given its security analysis.

8.1 Summary of the Thesis

In this thesis, we have defined a family of private-key block ciphers referred to as unbalanced CAST ciphers which employ the same structure of S-box and round function as the original CAST cipher but require a variable amount of memory depending on the chosen parameters. Furthermore, we select a set of typical parameters and investigate the security of the ciphers with respect of differential and linear cryptanalysis.

Although the unbalanced CAST ciphers constructed with 4×32 S-boxes require the least memory, a 4×32 S-box has at most 2^4 different 32-bit output vectors and

any linear combinations of such 2^4 vectors can not produce all 2^{32} different 32-bit vectors. Furthermore, a 4×32 S-box should have some affine boolean functions since its output bit size is much larger than its input bit size. Therefore, such S-boxes should be avoided because the potential risk to differential and linear cryptanalysis.

The unbalanced CAST ciphers with one or two 8×16 S-boxes are not as efficient as the unbalanced CAST ciphers configured by 8×32 S-boxes with respect to differential and linear cryptanalysis since the outputs of the round functions constructed with 8×16 S-boxes only influence half of the left half block bits. Also, it is very likely for the ciphers that iterative characteristics can be found with a high differential probability.

The 8×32 S-box, which is used by the original CAST cipher, seems to be most appropriate for the unbalanced CAST ciphers. The unbalanced CAST ciphers with one and two 8×32 S-boxes require only $1/4$ and $1/2$ the memory of the original CAST cipher, respectively. The result of analysis shows that the 48-round unbalanced CAST cipher with one 8×32 S-box and the 24-round unbalanced CAST cipher with two 8×32 S-boxes, which are equivalent to the 12-round original CAST cipher in efficiency, are resistant to both differential and linear cryptanalysis. We have also examined the unbalanced CAST ciphers from the perspective of information theory. As a result, the maximum static and dynamic input-output bit information leakages for the N/l-round unbalanced CAST ciphers constructed by the 8×32 S-boxes are much smaller than the corresponding ones for the 2-round DES.

The unbalanced CAST cipher with one 8×56 S-box is suitable for 64-bit proces-

sor implementations and a 40-round cipher is resistant to linear cryptanalysis. The differential attack methods studied in this thesis can not be applied to the cipher.

8.2 Suggestions for Further Research

The design and analysis of unbalanced CAST ciphers has not been completed. There are many paths that may be pursued to extend the analyses considered in this thesis.

Throughout this research, we have not investigated the design of a key schedule for unbalanced CAST ciphers. This issue should be explored further. A secure and efficient key schedule algorithm should be consequently proposed allowing for a variable-length cryptographic key.

Another interesting extension of the research presented in this thesis is to apply the attack based on non-surjective round functions [27] to unbalanced CAST ciphers. Since the round functions of all unbalanced CAST ciphers are non-surjective or non-uniform, this might be a new weakness for the ciphers, especially when the number of rounds is reduced to improve the encryption/decryption speed. Further research in this area is strongly encouraged.

References

[1] C. M. Adams, *Designing DES-Like Ciphers with Guaranteed Resistance to Differential and Linear Attacks*, Workshop on Selected Areas in Cryptography (SAC '95), pp. 133-144, Carleton University, Ottawa, Ontario, Canada, May 1995.

[2] C. M. Adams and S. E. Tavares, *Designing S-Boxes for Ciphers Resistant to Differential Cryptanalysis*, Proceedings of the 3rd Symposium on State and Progress of Research in Cryptography, pp. 181-190, Rome, Italy, February 1993.

[3] E. Biham and A. Shamir, *Differential Cryptanalysis of DES-Like Cryptosystems*, Journal of Cryptology, vol. 4, no. 1, pp. 3-72, 1991.

[4] E. Biham and A. Shamir, *Differential Cryptanalysis of FEAL and N-Hash*, Advances in Cryptology, Proceedings of EUROCRYPT '91, pp. 1-16, Springer-Verlag, 1991.

[5] E. Biham and A. Shamir, *Differential Cryptanalysis of Snefru, Khafre, REDOC-II, LOKI, and Lucifer*, Advances in Cryptology, Proceedings of CRYPTO '91, pp. 156-171, Springer-Verlag, 1991.

[6] L. Brown, M. Kwan, J. Pieprzyk and J. Seberry, *Improving Resistance to Differential Cryptanalysis and the Redesign of LOKI*, Advances in Cryptology, Proceedings of ASIACRYPT '91, pp. 36-50, Springer-Verlag, 1991.

[7] L. Brown, J. Pieprzyk and J. Seberry, *LOKI - A Cryptographic Primitive for Authentication and Secrecy Applications*, Advances in Cryptology, Proceedings of AUSCRYPT '90, pp. 229-236, Springer-Verlag, 1990.

[8] L. Brown and J. Seberry, *On the Design of Permutation P in DES Type Cryptosystems*, Advances in Cryptology, Proceedings of EUROCRYPT '89, pp. 696-705, Springer-Verlag, 1989.

[9] C. Chatfield, *Statistics for Technology*, John Wiley & Sons, 1979.

[10] M. H. Dawson and S. E. Tavares, *An Expanded Set of S-Box Design Criteria Based on Information Theory and its Relation to Differential-Like Attacks*, Advances in Cryptology, Proceedings of EUROCRYPT '91, pp. 352-367, Springer-Verlag, 1991.

[11] H. Feistel, *Cryptography and Computer Privacy*, Scientific American, vol. 228, no. 5, pp. 15-23, 1973.

[12] H. Feistel, W. A. Notz and J. L. Smith, *Some Cryptographic Techniques for Machine-to-Machine Data Communications*, Proceedings of the IEEE, vol. 63, no. 11, pp. 1545-1554, 1975.

[13] R. Forré, *Methods and Instruments for Designing S-Boxes*, Journal of Cryptology, vol. 2, pp. 115-130, 1990.

[14] H. M. Heys and S. E. Tavares, *On the Security of the CAST Encryption Algorithm*, Canadian Conference on Electrical and Computer Engineering, pp. 332-335, Halifax, Nova Scotia, Canada, September 1994.

[15] Burton S. Kaliski Jr. and Yiqun Lisa Yin, *On Differential and Linear Cryptanalysis of the RC5 Encryption Algorithm*, Advances in Cryptology, Proceedings of CRYPTO '95, pp. 171-184, Springer-Verlag, 1995.

[16] J. B. Kam and G. I. Davida, *Structured Design of Substitution-Permutation Encryption Networks*, IEEE Transactions on Computers, vol. 28, no. 10, pp. 747-753, 1979.

[17] X. Lai and J. Massey, *A Proposal for a New Block Encryption Standard*, Advances in Cryptology, Proceedings of EUROCRYPT '90, pp. 389-404, Springer-Verlag, 1990.

[18] J. Lee, H. M. Heys and S. E. Tavares, *On the Resistance of the CAST Encryption Algorithm to Differential Cryptanalysis*, Workshop on Selected Areas in Cryptography (SAC '95), pp. 107-120, Carleton University, Ottawa, Ontario, Canada, May 1995.

[19] M. Matsui, *Linear Cryptanalysis Method for DES Cipher*, Advances in Cryptology, Proceedings of EUROCRYPT '93, pp. 386-397, Springer-Verlag, 1993.

[20] W. Meier and O. Staffelbach, *Nonlinearity Criteria for Cryptographic Functions*, Advances in Cryptology, Proceedings of EUROCRYPT '89, pp. 549-562, Springer-Verlag, 1989.

[21] R. C. Merkle, *Fast Software Encryption Functions*, Advances in Cryptology, Proceedings of CRYPTO '90, pp. 627-638, Springer-Verlag, 1990.

[22] S. Mister and C. M. Adams, *Practical S-Box Design*, Workshop on Selected Areas in Cryptography (SAC '96), pp. 61-76, Queen's University, Kingston, Ontario, Canada, August 1996.

[23] S. Miyaguchi, *The FEAL Cipher Family*, Advances in Cryptology, Proceegings of CRYPTO '90, pp. 627-638, Springer-Verlag, 1990.

[24] National Bureau of Standards, *Data Encryption Standard*, Federal Information Processing Standard Publication 46, 1977.

[25] K. Nyberg, *Perfect Nonlinear S-Boxes*, Advances in Cryptology, Proceedings of EUROCRYPT '91, pp. 378-386, Springer-Verlag, 1991.

[26] K. Ohta and K. Aoki, *Linear Cryptanalysis of the Fast Data Encipherment Algorithm*, Advances in Cryptology, Proceedings of CRYPTO '94, pp. 12-16, Springer-Verlag, 1994.

[27] V. Rijmen and B. Preneel, *On Weakness of Non-Surjective Round Functions*, Workshop on Selected Areas in Cryptography (SAC '95), pp. 100-106, Carleton University, Ottawa, Ontario, Canada, May 1995.

[28] R. Rivest, *The RC-5 Encryption Algorithm*. Proceedings of the Second International Workshop on Fast Software Encryption, pp. 86-96. Springer-Verlag, 1995.

[29] O. S. Rothaus, *On Bent Functions*, Journal of Combinatorial Theory, vol. 20, no. A, pp. 300-305, 1976.

[30] B. Schneier, *The Blowfish Encryption Algorithm*, Proceedings of the Cambridge Security Workshop on Fast Software Encryption, pp. 191-204, U. K., December 9-11, 1993.

[31] C. E. Shannon, *Communication Theory of Secrecy Systems*, Bell System Technical Journal, vol. 28, pp. 656-715, 1949.

[32] A. Shimizu and S. Miyaguchi, *Fast Data Encipherment Algorithm FEAL*, Advances in Cryptology, Proceedings of EUROCRYPT '87, pp. 267-278, Springer-Verlag, 1987.

[33] M. Sivabalan, S. E. Tavares and L. E. Peppard, *On the Design of SP Networks from an Information Theoretic Point of View*, Advances of Cryptology, Proceedings of CRYPTO '92, pp. 260-279, Springer-Verlag, 1992.

[34] A. F. Webster and S. E. Tavares, *On the Design of S-Boxes*, Advances in Cryptology, Proceedings of CRYPTO '85, pp. 523-534, Springer-Verlag, 1985.

[35] M. J. Wiener, *Efficient DES Key Search*, presented at CRYPTO '93, Santa Barbara, California, USA, August 1993.

[36] A. M. Youssef, S. E. Tavares, S. Mister and C. M. Adams, *Linear Approximation of Injective S-Boxes*, IEE Electronics Letters, vol. 31, no. 25, pp. 2168-2169, 1995.

[37] M. Zhang, S. E. Tavares and L. L. Campbell, *Information Leakage of Boolean Functions and its Relationship to Other Cryptographic Criteria*, Proceedings of 2nd ACM Conference on Computer and Communications Security, pp. 156-165, Fairfax, Virginia, USA, November 1993.